

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:13:34 ; Search time 163.5 Seconds
(without alignments)
4056.844 Million:cell updates/sec

Title: US-10-021-571-2

Perfect score: 8853

Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSESTGFGBERESIL 1715

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8853	100.0	1715	5	AAE25143 Rat ARMS
2	8853	100.0	1715	7	Adc06849 Rat Kidin
3	8639.5	97.6	1762	5	Aau96841 Rat Kidin
4	8639.5	97.6	1762	7	Adc06848 Rat Kidin
5	8526	96.3	1763	5	Aau80244 Rat Kidin
6	8266	93.4	1715	4	Aam39025 Human pol
7	8266	93.4	1715	5	Aam38993 Human pol
8	8266	93.4	1715	5	AAE25144 Human ARM
9	8266	93.4	1715	5	Aau96840 Human kid
10	8266	93.4	1715	6	AAE32128 Human cyt
11	8237	93.0	1753	6	Ada09888 Human rec
12	8212	92.8	1771	7	Adc06847 Human Kid
13	5747	64.9	1142	7	Adc06850 Kidins 22
14	5730.5	64.7	1184	7	Adc06846 Human Kid
15	2961.5	33.5	705	4	Aag75604 Human col
16	2782.5	31.4	551	4	AAE01035 Human dea
17	2408	27.2	543	4	AAE95191 Human pro
18	2376.5	26.8	1498	4	ABE64857 Drosophil
19	2376.5	26.8	1498	5	AAE25146 Fruit fly
20	2260	25.5	513	4	AAE03645 Human ext
21	2150	24.3	624	4	ABG08697 Novel hum
22	1848.5	20.9	1398	5	AAE25145 Worm ARMS
23	662	7.5	131	5	ADK36597 Novel hum
24	614	6.9	120	4	AAE40811 Human pol
25	599.5	6.8	129	4	AAE94785 Human pro

26	558	6.3	1881	7	ADD47763 Human Pro
27	553	6.2	1880	8	ADR90358 Full leng
28	553	6.2	3897	8	ABM83919 Human dia
29	553	6.2	3924	8	ABM83916 Human dia
30	553	6.2	3955	8	ABM83915 Human dia
31	553	6.2	3985	8	ABM83914 Human dia
32	553	6.2	4016	8	ABM83912 Human dia
33	553	6.2	4072	8	ABM83913 Human dia
34	549.5	6.2	4082	8	ABM83911 Human dia
35	549	6.2	109	3	AAE56277 Human sec
36	542.5	6.1	1955	8	ABM83917 Human dia
37	542	6.1	1954	8	ABM83918 Human dia
38	538.5	6.1	1724	5	AAO20513 Protein o
39	536	6.1	1839	7	ADD27862 Human bra
40	534.5	6.0	1872	4	AAE79160 Human pro
41	531.5	6.0	986	4	ABG31785 Human ank
42	531.5	6.0	1139	7	ADC39122 Novel hum
43	531.5	6.0	1250	6	AAE33667 Human str
44	529	6.0	2622	7	ADE55508 Rat Prote
45	527.5	6.0	657	7	ADB64375 Human pro

ALIGNMENTS

RESULT 1

AAE25143
ID AAE25143 standard; protein; 1715 AA.

XX AC AAE25143;

DT 30-OCT-2002 (first entry)

DE Rat ARMS protein.

XX

XX Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
XX diagnostic; imaging; rat.

OS Rattus norvegicus.

XX	Key	Location/Qualifiers
FT	Region	17..390
FT		/note= "Ankyrin repeat"
FT	Region	354..493
FT		/note= "N-terminal region"
FT	Domain	496..518
FT		/note= "Transmembrane domain"
FT	Domain	525..546
FT		/note= "Transmembrane domain"
FT	Region	573..638
FT		/note= "Loop between TM2 and TM3"
FT	Domain	661..680
FT		/note= "Transmembrane domain"
FT	Domain	688..710
FT		/note= "Transmembrane domain"
FT	Region	749..854
FT		/note= "Carboxy tail"
FT	Region	940..1060
FT		/note= "Carboxy tail"
FT	Region	1081..1093
FT		/note= "Polyproline stretch"
FT	Domain	1151..1221
FT		/note= "SAM domain"
FT	Binding-site	1713..1715
FT		/note= "PDZ binding motif"

WO200250273-A2.

27-JUN-2002.

19-DEC-2001; 2001WO-US048603.

XX

RESULT 2
ADC06849
ID ADC06849 standard; protein; 1715 AA.
XX
AC ADC06849;
XX
DT 18-DEC-2003 (first entry)
XX
DZ Rat Kidins220Pc protein AF313464.
XX
KW cytostatic; prostate cancer; breast; gene therapy; transgenic; rat;
KW Kidins220Pc; kinase D-interacting substrate of 220kDa.
XX
OS Rattus sp.
XX
FN WO2003064599-A2.
XX
PD 07-AUG-2003.
XX
PF 24-JAN-2003; 2003WO-US001943.
XX
PR 25-JAN-2002; 2002US-00054935.
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Pan W, Shu Y;
XX
DR WPI; 2003-679495/64.
XX
PT New isolated polynucleotide related to cancer genes, useful for
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
PT or treating cancers, e.g. breast and prostate cancers.
XX
XX Disclosure; Fig 20; 128pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC differentially-regulated mammalian cancer gene. The polynucleotides of
CC the invention demonstrate cytostatic activity and are differentially
CC expressed in prostate cancer. The polynucleotide, polypeptides and
CC methods of the invention may be useful for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating cancers,
CC particularly breast and prostate cancers. Furthermore, the invention may
CC be utilised during gene therapy procedures or in the production of
CC transgenic animals. The current sequence is that of the prostate cancer-
CC related protein of the invention. The current sequence is not fully
CC legible within the specification; the complete sequence was obtained from
CC GenBank.
XX
SQ Sequence 1715 AA;

Query Match 100.0%; Score 8853; DB 7; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSVLISQSVINVEENIPALKALLECKVDNERNECCQTPMLAABQGNVEIVKELKN 60
Db 1 MSVLISQSVINVEENIPALKALLECKVDNERNECCQTPMLAABQGNVEIVKELKN 60

Qy 61 GANCNLEDLDNWTALISASKEGHIHVEELLKSGASLEHRDGGWTALMWACYKGRDGV 120
Db 61 GANCNLEDLDNWTALISASKEGHIHVEELLKSGASLEHRDGGWTALMWACYKGRDGV 120

Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIHLLLQNGAKVNCSDKYGTTPPLWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADIHLLLQNGAKVNCSDKYGTTPPLWAAR 180

Qy 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYYTOSVKEILLKRNPNVNLTDKGNTA 240
Db 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYYTOSVKEILLKRNPNVNLTDKGNTA 240

Qy 241 LMIASKEGHIHVEODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMIASKEGHIHVEODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300

Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATGRIEIVVEVLLDKGA 360
Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATGRIEIVVEVLLDKGA 360

Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQSKIL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQSKIL 420

Qy 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSPTMQPPICVGLYAGWGSGKSLKKL 480
Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSPTMQPPICVGLYAGWGSGKSLKKL 480

Qy 481 EDEMTFAGQOTEPLFQFSWLIIVFLTLCCGGLGVPAFPVDTNLAIASISFLAIYIF 540
Db 481 EDEMTFAGQOTEPLFQFSWLIIVFLTLCCGGLGVPAFPVDTNLAIASISFLAIYIF 540

Qy 541 FIVYFGRRREGESNNAWALSTRLARHIGYLELLFKLMFVNPPELPQTTKALPVRFLF 600
Db 541 FIVYFGRRREGESNNAWALSTRLARHIGYLELLFKLMFVNPPELPQTTKALPVRFLF 600

Qy 601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660
Db 601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660

Qy 661 FVIFLFIVGCIAGITLLAIFRVDPKHVTNAILISIASVVGGLAFVLNCRTWQVLDLIL 720
Db 661 FVIFLFIVGCIAGITLLAIFRVDPKHVTNAILISIASVVGGLAFVLNCRTWQVLDLIL 720

Qy 721 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMARKTIDSFQNTQRLVVIDGLDAC 780
Db 721 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMARKTIDSFQNTQRLVVIDGLDAC 780

Qy 781 EQDKVLOMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLDRDSNINGHDYMRNVH 840
Db 781 EQDKVLOMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLDRDSNINGHDYMRNVH 840

Qy 841 LPVFLNRSGLSNARKELVTSATNGDITCSDTGTQEDTDREVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNRSGLSNARKELVTSATNGDITCSDTGTQEDTDREVSONSLGEMTKLGSKTALN 900

Qy 901 RRDYRRRQMQRTITRQMSFDLTLLVTEDFWFSDISPQTMERLLNIVSVTGRLLRANQIT 960
Db 901 RRDYRRRQMQRTITRQMSFDLTLLVTEDFWFSDISPQTMERLLNIVSVTGRLLRANQIT 960

Qy 961 FNWDLASWINLTQWPYRTSWLILYEETEGLPDQMTLKTMYERISKNIPPTKDVPELL 1020
Db 961 FNWDLASWINLTQWPYRTSWLILYEETEGLPDQMTLKTMYERISKNIPPTKDVPELL 1020

Qy 1021 EIDGDIENFEVFLSSRTPVLVARDVKTFPLCTVNLDPKIREIIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIENFEVFLSSRTPVLVARDVKTFPLCTVNLDPKIREIIADVRAAREQINIGGLAY 1080

Qy 1081 PPLPHLGPPPPPGSYQSPASVCSASFNGPFGVSPQPHSSYVSGLSGQPHFPYNA 1140
Db 1081 PPLPHLGPPPPPGSYQSPASVCSASFNGPFGVSPQPHSSYVSGLSGQPHFPYNA 1140

Qy 1141 AVPATGSSLLSSMTVDVVCCKLQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200
Db 1141 AVPATGSSLLSSMTVDVVCCKLQIEGLDQNMMPQYCTTIKKANINGRVLSQCNIDELK 1200

Qy 1201 KEMAMNFGDWHLFRSMVLEMRSVESQVVPDPREFLNENSSAPVPHGESARSSHTLEPIT 1260
Db 1201 KEMAMNFGDWHLFRSMVLEMRSVESQVVPDPREFLNENSSAPVPHGESARSSHTLEPIT 1260

Qy 1261 ELSQOTPYTLNFSPEELNTLGLDEGAPRHSNLSMQSOTRRTPSLSSLSQDSSEISKLT 1320

Db 1261 ELSSQTPVTLNFSFEELNLTGLDEGAPRHNSLWSQSTRTPSLSLNSQSSIEISKLT 1320
Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGGSIHSTLEQERG 1380
Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGGSIHSTLEQERG 1380
Qy 1381 KEGELKQEDGKRSFLMKRGDIVDYSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGKRSFLMKRGDIVDYSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLQTDLKLKGGGLRYQKLPSEDESGTRGVQITPHCSKMIRTKRLKAKQRECA SP 1500
Db 1441 SERPSLQTDLKLKGGGLRYQKLPSEDESGTRGVQITPHCSKMIRTKRLKAKQRECA SP 1500
Qy 1501 QEHSAPERTFIRKAYEYLSALLDKKSDSGVRSNNESSPHSHNEAADDSOLEKANLI 1560
Db 1501 QEHSAPERTFIRKAYEYLSALLDKKSDSGVRSNNESSPHSHNEAADDSOLEKANLI 1560
Qy 1561 ELEDGHSGRGMPHSLGSLQDPTIARMSICSEDKKSPSCSLIASSPESWPACOKAYN 1620
Db 1561 ELEDGHSGRGMPHSLGSLQDPTIARMSICSEDKKSPSCSLIASSPESWPACOKAYN 1620
Qy 1621 LNRTPTVTLNNNTAPTNRANQNFDEIGRETSQVILRPGSPNPTAVQENLKSMAHK 1680
Db 1621 LNRTPTVTLNNNTAPTNRANQNFDEIGRETSQVILRPGSPNPTAVQENLKSMAHK 1680
Qy 1681 RSQRSSYTRLSDKASELHAASSESTGFGEREESIL 1715
Db 1681 RSQRSSYTRLSDKASELHAASSESTGFGEREESIL 1715

RESULT 3

AAU96841
ID AAU96841 standard; protein; 1762 AA.

XX AAU96841;

DT 09-SEP-2004 (revised)

DT 30-JUL-2002 (first entry)

XX Rat kidins220 protein from Genbank AAG35185.

XX Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
XX neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
XX neurodegenerative disease; glioblastoma multiforme; prostate cancer.

OS Rattus sp.
OS Unidentified.

XX W0200220786-A2.

XX 14-MAR-2002.

XX 06-SEP-2001; 2001WO-GB003977.

XX 06-SEP-2000; 2000US-0230449P.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

XX Schiavo G, Iglesias T;

XX WPI; 2002-371879/40.

XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,
XX useful for identifying modulators useful in treating cancer and
XX neurodegenerative diseases.

XX Disclosure; Fig 18; 228pp; English.

XX The invention relates to polypeptide comprising a kinase D interacting
XX substrate of 220 kDa (Kidins220) from rat and human or their fragment,
XX variant or fusion provided that the protein is not the polypeptide

CC encoded by GenBank Accession No. AB033076. Also included are a
CC polynucleotide encoding Kidins220 provided that polynucleotide is not any
CC one of the clones corresponding to the 61 GenBank Accession Nos. given in
CC the specification; an expression/replicable vector comprising the
CC polynucleotide; a recombinant host cell containing the polynucleotide or
CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)
CC ; an agent (A1) which modulates activity of Kidins220, protein kinase D
CC (PKD) or interaction between PKD and Kidins220; a mutant animal
CC transgenic for kidins220; the use of an agent capable of detecting the
CC expression of kidins220 gene in manufacture of a diagnostic reagent for
CC diagnosing or prognosing cancer or for monitoring the progression of
CC cancer in a patient; a compound comprising a moiety that selectively
CC binds to Kidins220 or its variant and another moiety; and a nucleic acid
CC molecule encoding the compound. The antibody is useful for modulating the
CC activity of Kidins220, where the antibody prevents a region of Kidins220
CC interacting from another protein, or reduces the ability of Kidins220 to
CC bind to adenosine triphosphate (ATP). The kidins220 polynucleotide is
CC useful for identifying an agent which modulates the promoter activity of
CC the polynucleotide, and Kidins220 is useful for identifying an agent
CC which modulates activity of Kidins220, protein kinase D (PKD) and the
CC interaction between PKD with Kidins220. Kidins220 or its polynucleotide
CC is useful in medicine, e.g. using gene therapy or for diagnosing or
CC monitoring progression of cancer in a patient. The agent is useful in
CC manufacture of medicament, for use in treatment of neurodegenerative
CC disease, and the agent which inhibits function of kidins220 gene or its
CC product is useful for treating cancer which is glioblastoma multiforme or
CC prostate cancer. The compound is useful for imaging cancer in an
CC individual, and for diagnosing or prognosing, and also for treating an
CC individual. The present sequence represents Rat kidins220 protein from
CC Genbank accession number AAG35185

CC Revised record issued on 09-SEP-2004 : Correction to Organism field

XX Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 5; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;
Qy 1 MSVLISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGVVEIVKLLKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMLAAEQGVVEIVKLLKN 60
Qy 61 GANCNLEDJNWTALISASKEGHIHVIBELLKSGASLEHRDMGCMWTALMWACVKGRTDVV 120
Db 61 GANCNLEDJNWTALISASKEGHIHVIBELLKSGASLEHRDMGCMWTALMWACVKGRTDVV 120
Qy 121 ELLLSHGANSVVTGLQYSVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLYWAAR 180
Db 121 ELLLSHGANSVVTGL--YSVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLYWAAR 179
Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTSQVKEILKRNPNVNLTKDGNNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTSQVKEILKRNPNVNLTKDGNNTA 239
Qy 241 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIDIRG 300
Db 240 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIDIRG 299
Qy 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLRPNKAGETPNIDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLRPNKAGETPNIDCSHOKSIL 419
Qy 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSKFLKKL 480
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGSKFLKKL 479
Qy 481 EDESKTFAQQOETPLFQFSLWLVFLTLCCGGLGVFAFPVDTNLAIAISLSFLAIYIF 540

Db 480 EDEMTFAGQQTPELPQPSWLIIVFUTLLCGGLGLVFPFTVDTNLAIASISFLALIYIF 539
Qy 541 FIVIFYGGRREGESNMAWALSTRLARHIGYLELLFKLMFVNPPPELQPTTKALPVRFLF 600
Db 540 FIVIFYGGRREGESNMAWALSTRLARHIGYLELLFKLMFVNPPPELQPTTKALPVRFLF 599
Qy 601 TDYRLSSVGGETSIAEMIALSDACEREFGLATRLFRVFRTERESQKKKKWKTCCCLPS 660
Db 600 TDYRLSSVGGETSIAEMIALSDACEREFGLATRLFRVFRTERESQKKKKWKTCCCLPS 659
Qy 661 FVIFLFIIVGCIAGITLAIIPRVDPKHLTVNAIILSIASVVGCLAVLNCRTWQVLDL 720
Db 660 FVIFLFIIVGCIAGITLAIIPRVDPKHLTVNAIILSIASVVGCLAVLNCRTWQVLDL 719
Qy 721 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLVVIIDGLDAC 780
Db 720 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLVVIIDGLDAC 779
Qy 781 EQDKVLQMLDTRVRLFSKGPPIAIPASDPHIIKAINQNLSVLRDSNINGHDYMRNIVH 840
Db 780 EQDKVLQMLDTRVRLFSKGPPIAIPASDPHIIKAINQNLSVLRDSNINGHDYMRNIVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSONSLGEMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSONSLGEMTKLGSKTALN 899
Qy 901 RRDYRRQMQRTITRQMSFDLTKLLVTEDEFSDISPTQMRLLNIVSVTGRLLRANQIT 960
Db 900 RRDYRRQMQRTITRQMSFDLTKLLVTEDEFSDISPTQMRLLNIVSVTGRLLRANQIT 959
Qy 961 FNRDLASWINLTQWVPYRTSWLILYLBETEGLPDQMTLTKMYERISKNIPPTTKDVRPLL 1020
Db 960 FNRDLASWINLTQWVPYRTSWLILYLBETEGLPDQMTLTKMYERISKNIPPTTKDVRPLL 1019
Qy 1021 EIDGDIRNFVFLSRTPLVARDKTFLPCTVNLDPKRLIADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSRTPLVARDKTFLPCTVNLDPKRLIADVRAAREQINIGGLAY 1079
Qy 1081 PPLPLHEGPPPPSGYSPASVCSASFPNGVSPQPHSSYVSGSLGSPQHPFYNR- 1139
Db 1080 PPLPLHEGPPPPSGYSPASVCSASFPNGVSPQPHSSYVSGSLGSPQHPFYNRP 1139
Qy 1140 -----AAVPATGSSILL 1151
Db 1140 PFAPLYTPRYPGSQHLSRSVKTSPLRDQNGPLPCDSGFNKQRAVPAATGSSILL 1199
Qy 1152 SSMTVDVCEKRLQIEGLDQNMPOYCTTIKKNINGRVLQCNIDELKEMANFGDWH 1211
Db 1200 SSMTVDVCEKRLQIEGLDQNMPOYCTTIKKNINGRVLQCNIDELKEMANFGDWH 1259
Qy 1212 LFRSMVLEMRVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPTELSSQTPYTLN 1271
Db 1260 LFRSMVLEMRVESQVVPEDPRFLNENSSAPVPHGESARRSSHTELPTELSSQTPYTLN 1319
Qy 1272 PSPELNLGLDEGAPRHSNLWSQSTRTPSLSSLNSQDSSIBISKLTDKVQAEYRDAY 1331
Db 1320 PSPELNLGLDEGAPRHSNLWSQSTRTPSLSSLNSQDSSIBISKLTDKVQAEYRDAY 1379
Qy 1332 REYTAQMSQLEGGTSSITSGRSSPHSTYIYGQSSSGSIHSTLEQERKEGELKQEDGR 1391
Db 1380 REYTAQMSQLEGGTSSITSGRSSPHSTYIYGQSSSGSIHSTLEQERKEGELKQEDGR 1439
Qy 1392 KSFLMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSLAPGKKSSRPSLFTD 1451
Db 1440 KSFLMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSLAPGKKSSRPSLFTD 1499
Qy 1452 KLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIITKRLKAK-QRECAAPQESAEPIRT 1510
Db 1500 KLKGGGLRYQKLPSDEDESGTGRVQITPHCSKMIITKRLKAK-QRECAAPQESAEPIRT 1557
Qy 1511 PIKAKYLSDALDKQSSDSGVRSNENSSPHSLHNEAADSQLEKANLIELEDEGHSGK 1570
Db 1558 PIKAKYLSDALDKQSSDSGVRSNENSSPHSLHNEAADSQLEKANLIELEDEGHSGK 1617

Qy 1571 RGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTSTVTL 1630
Db 1618 RGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTSTVTL 1677
Qy 1631 NNNTAPTNRANQNDFEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQRSSYTRL 1690
Db 1678 NNNTAPTNRANQNDFEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQRSSYTRL 1737
Qy 1691 SKDASELHAASSESTGFGEERESIL 1715
Db 1738 SKDASELHAASSESTGFGEERESIL 1762

RESULT 4

ADC06848
ID ADC06848 standard; protein; 1762 AA.

XX
AC ADC06848;

XX
DT 18-DEC-2003 (first entry)

XX
DE Rat Kidins220Pc protein AF239045.

XX
KW cytotatic; prostate cancer; breast; gene therapy; transgenic; rat;
Kidins220Pc; kinase D-interacting substrate of 220KDa.

XX
OS Rattus sp.

XX
PN WO2003064599-A2.

XX
PD 07-AUG-2003.

XX
PF 24-JAN-2003; 2003WO-US001943.

XX
PR 25-JAN-2002; 2002US-00054935.

PR 14-FEB-2002; 2002US-0356130P.

PR 22-MAR-2002; 2002US-00102946.

PR 08-APR-2002; 2002US-00117229.

PR 14-MAY-2002; 2002US-00144198.

PR 19-JUL-2002; 2002US-00197824.

XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX
PI Sun Z, Li X, Jay G, Kovacs KP, Fan W, Shu Y;

XX
DR WPI; 2003-679495/64.

XX
PT New isolated polynucleotide related to cancer genes, useful for
detecting, diagnosing, staging, monitoring, prognosticating, preventing

XX
PS or treating cancers, e.g. breast and prostate cancers.

XX
PS Disclosure; Fig 20; 128pp; English.

XX
CC The invention relates to a novel isolated polynucleotide comprising a
differentially-regulated mammalian cancer gene. The polynucleotides of
the invention demonstrate cytostatic activity and are differentially
expressed in prostate cancer. The polynucleotide, polypeptides and
methods of the invention may be useful for detecting, diagnosing,
staging, monitoring, prognosticating, preventing or treating cancers,
particularly breast and prostate cancers. Furthermore, the invention may
be utilized during gene therapy procedures or in the production of
transgenic animals. The current sequence is that of the prostate cancer-
related protein of the invention. The current sequence is not fully
legible within the specification; the complete sequence was obtained from
GenBank.

XX
SQ Sequence 1762 AA;

Query Match 97.6%; Score 8639.5; DB 7; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;

QY 1 MSLVLSQSVINYVEENIPALKALLBKCKDVDERNECGOTPLMLAAEQGNVEIVKELLKN 60
DB 1 MSLVLSQSVINYVEENIPALKALLBKCKDVDERNECGOTPLMLAAEQGNVEIVKELLKN 60
QY 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACVKGRTDVV 120
DB 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGGWTALMWACVKGRTDVV 120
QY 121 ELLLSHGANSVTGLOYSVYPIIWAAGRGHADIHVHLLQNGAKVNSDKYGTTPLYWAAR 180
DB 121 ELLLSHGANSVTGLOYSVYPIIWAAGRGHADIHVHLLQNGAKVNSDKYGTTPLYWAAR 179
QY 181 KGHLECVKHLANGADVDOGANSMTALIVAVKGGYTTQSVKEILKRNPNVNLTDKDGNTA 240
DB 180 KGHLECVKHLANGADVDOGANSMTALIVAVKGGYTTQSVKEILKRNPNVNLTDKDGNTA 239
QY 241 LMTASKEGHIIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
DB 240 LMTASKEGHIIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVERGNATWVRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 360
DB 300 QDNKTALYWAVERGNATWVRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 359
QY 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420
DB 360 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 419
QY 421 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFKLLK 480
DB 420 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFKLLK 479
QY 481 EDEWKTFAQOOTEPLFOFSLIIVFLTLILCGGLVFAFVDTNLAIAISFLAIYIF 540
DB 480 EDEWKTFAQOOTEPLFOFSLIIVFLTLILCGGLVFAFVDTNLAIAISFLAIYIF 539
QY 541 FIVYFGREGESNNWAWLSRFLARHIGVLELLFKLMFVNPPPELPEQTTKALPVRF 600
DB 540 FIVYFGREGESNNWAWLSRFLARHIGVLELLFKLMFVNPPPELPEQTTKALPVRF 599
QY 601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKKTCLPS 660
DB 600 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQKKWKKTCLPS 659
QY 661 FVIFLFTVIGCIAGITTLAIIFRVDPKHLTVNAILISIASVVGAFVNLCTWQVLDLSL 720
DB 660 FVIFLFTVIGCIAGITTLAIIFRVDPKHLTVNAILISIASVVGAFVNLCTWQVLDLSL 719
QY 721 NSQRKRLHSAASKLHLKSGFMKVKCEVELMARMAKTIDSFQNOTRLVWIIDGLDAC 780
DB 720 NSQRKRLHSAASKLHLKSGFMKVKCEVELMARMAKTIDSFQNOTRLVWIIDGLDAC 779
QY 781 EQDKVLQMLDTRVFLFSKGPPIAIFASDPHIIKAINONLSNVLSDNSNINGHYMNRV 840
DB 780 EQDKVLQMLDTRVFLFSKGPPIAIFASDPHIIKAINONLSNVLSDNSNINGHYMNRV 839
QY 841 LPVFLNSRGLSNARKFLVTSNMGDITCSDTTGTQEDTDRVSONSIGEMTKLGSKTALN 900
DB 840 LPVFLNSRGLSNARKFLVTSNMGDITCSDTTGTQEDTDRVSONSIGEMTKLGSKTALN 899
QY 901 RRDYTRRQWORTITRQMSFDLTKLLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIT 960
DB 900 RRDYTRRQWORTITRQMSFDLTKLLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIT 959
QY 961 FNWDRLASWINLTQWVPYRTSWLILYEETEGLPDQMTLKTMYERIKNPTTKDVEPFL 1020
DB 960 FNWDRLASWINLTQWVPYRTSWLILYEETEGLPDQMTLKTMYERIKNPTTKDVEPFL 1019
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKTLFCTVNLDPKLEIADVRAAEQINIGGLAY 1080
DB 1020 EIDGDIRNFVFLSSRTPVLVARDVKTLFCTVNLDPKLEIADVRAAEQINIGGLAY 1079
QY 1081 PPLPLHEGPPRPPSGYSPASVCSASFNGFPFGVSPQPHSSYYSGLSQPHFYNR - 1139

DB 1080 PPLPLHEGPPRPPSGYSPASVCSASFNGFPFGVSPQPHSSYYSGLSQPHFYNR 1139
QY 1140 -----AAVPATGSSLL 1151
DB 1140 PFAPLYLTPRYPPGSGQHLISRSSVKTSLPRDNNGLPCDSGFNQRQAAVPATGSSLL 1199
QY 1152 SSMVTVDVCEKURQIEGLDQNNMPOYCTTIKCANINGRVLSCNIDELKEMAMNFGDWH 1211
DB 1200 SSMVTVDVCEKURQIEGLDQNNMPOYCTTIKCANINGRVLSCNIDELKEMAMNFGDWH 1259
QY 1212 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTIN 1271
DB 1260 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTIN 1319
QY 1272 FSFEELNTLGLDEGAPRHSNLSWQOTRTPSLNSODSSIEISKLTDKQVQAEVRDAY 1331
DB 1320 FSFEELNTLGLDEGAPRHSNLSWQOTRTPSLNSODSSIEISKLTDKQVQAEVRDAY 1379
QY 1332 REVIAQMSOLEGTGSGSTISGRSSPHSTYIIGQSSGSGSIHSTLEQERKEGELKQEDGR 1391
DB 1380 REVIAQMSOLEGTGSGSTISGRSSPHSTYIIGQSSGSGSIHSTLEQERKEGELKQEDGR 1439
QY 1392 KSFLMKRGDVIDYSSGSGVTNEASPLDPITEDEKSDQSGSKLLPKKKSERPSLPQTDL 1451
DB 1440 KSFLMKRGDVIDYSSGSGVTNEASPLDPITEDEKSDQSGSKLLPKKKSERPSLPQTDL 1499
QY 1452 KLKGGGLRYOKLPSEDESGETGRVQITPHCSKMIRTKLAK - QRECAPSQEHSASPRT 1510
DB 1500 KLKGGGLRYOKLPSEDESGETGRVQITPHCSKMIRTKLAK - QRECAPSQEHSASPRT 1557
QY 1511 FIKAKYVLSDALLDKDDSDSGVRSNESPNSHLSHNEAADDSDQLEKANLIELEDEGHSGK 1570
DB 1558 FIKAKYVLSDALLDKDDSDSGVRSNESPNSHLSHNEAADDSDQLEKANLIELEDEGHSGK 1617
QY 1571 RGNPHSLSGLOPPIIARMSICSEDKKSPSECSLIASPEESWPACQKAYNLNRTPTVTIL 1630
DB 1618 RGNPHSLSGLOPPIIARMSICSEDKKSPSECSLIASPEESWPACQKAYNLNRTPTVTIL 1677
QY 1631 NNNTAPTNRANQNFDEIGRIETTSQVILRPGSPNPTAVQENLKSMAHKRSQSSVTRL 1690
DB 1678 NNNTAPTNRANQNFDEIGRIETTSQVILRPGSPNPTAVQENLKSMAHKRSQSSVTRL 1737
QY 1691 SKDASELHAASSESTGFGEERESIL 1715
DB 1738 SKDASELHAASSESTGFGEERESIL 1762
RESULT 5
AAU80244
ID AAU80244 standard; protein; 1763 AA.
XX AAU80244;
AC AAU80244;
XX XX
DT 09-SEP-2004 (revised)
DT 30-JUL-2002 (first entry)
XX XX
DE Rat kidins220 protein.
XX XX
KW Kidins220; kinase D interacting substrate of 22kDa; cytosstatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer.
OS Rattus sp.
OS Unidentified.
XX XX
PN WO200220786-A2.
XX XX
PD 14-MAR-2002.
XX XX
PF 06-SEP-2001; 2001WO-GB003977.
XX XX
PR 06-SEP-2000; 2000US-0230449P.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX Schiavo G, Iglesias T;
 XX WPI: 2002-371879/40.
 XX N-PSDB; ABK51214.
 XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,
 PT useful for identifying modulators useful in treating cancer and
 PT neurodegenerative diseases.
 XX Claim 33; Fig 1; 228pp; English.

The invention relates to polypeptide comprising a kinase D interacting substrate of 220 kDa (Kidins220) from rat and human or their fragment, variant or fusion provided that the protein is not the polypeptide encoded by GenBank Accession No. AB033076. Also included are a polynucleotide encoding Kidins220 provided that polynucleotide is not any one of the clones corresponding to the 61 GenBank Accession Nos. given in the specification; an expression/replicable vector comprising the polynucleotide; a recombinant host cell containing the polynucleotide or vector; an anti-Kidins220 antibody (used in the preparation of Kidins220) ; an agent (AI) which modulates activity of Kidins220, protein kinase D (PKD) or interaction between PKD and Kidins220; a mutant animal transgenic for Kidins220; the use of an agent capable of detecting the expression of Kidins220 gene in manufacture of a diagnostic reagent for diagnosing or prognosing cancer or for monitoring the progression of cancer in a patient; a compound comprising a moiety that selectively binds to Kidins220 or its variant and another moiety; and a nucleic acid molecule encoding the compound. The antibody is useful for modulating the activity of Kidins220, where the antibody prevents a region of Kidins220 interacting from another protein, or reduces the ability of Kidins220 to bind to adenosine triphosphate (ATP). The Kidins220 polynucleotide is useful for identifying an agent which modulates the promoter activity of the polynucleotide, and Kidins220 is useful for identifying an agent which modulates activity of Kidins220, protein kinase D (PKD) and the interaction between PKD with Kidins220. Kidins220 or its polynucleotide is useful in medicine, e.g. using gene therapy or for diagnosing or monitoring progression of cancer in a patient. The agent is useful in manufacture of medicament, for use in treatment of neurodegenerative disease, and the agent which inhibits function of Kidins220 gene or its product is useful for treating cancer which is glioblastoma multiforme or prostate cancer. The compound is useful for imaging cancer in an individual, and for diagnosing or prognosing, and also for treating an individual. The present sequence represents Rat kidins220

Revised record issued on 09-SEP-2004 : Correction to Organism field

Sequence 1763 AA;

Query Match 96.3%; Score 8526; DB 5; Length 1763;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1671; Conservative 10; Mismatches 32; Indels 52; Gaps 6;

Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECQTPMLAAEQGNVEIVKELKN 60
 Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECQTPMLAAEQGNVEIVKELKN 60
 Qy 61 GANCNLELDNWTALISAKEGHIIHVEELKSGASLEHRDMGGWTALMWAICYKGRDVV 120
 Db 61 GANCNLELDNWTALISAKEGHIIHVEELKSGASLEHRDMGGWTALMWAICYKGRDVV 120
 Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYCTTPLVWAAR 180
 Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYCTTPLVWAAR 180
 Qy 181 KGHLECVKHLAMGADVDEGANSTALIIVAKGGYTQSVKEILKRNPNVNLTKDGNTA 240
 Db 181 KGHLECVKHLAMGADVDEGANSTALIIVAKGGYTQSVKEILKRNPNVNLTKDGNTA 240
 Qy 241 LMIASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
 Db 241 LMIASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

Db 241 LMIASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
 Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGETPLIKATKWRNIEVVVELLDKGA 360
 Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGETPLIKATKWRNIEVVVELLDKGA 360
 Qy 361 KVSADVKKGDTPLHVAIRGRSRRRLAELLLRNPKOGRLLYRPNKAGETPNIDCSHQSKSL 420
 Db 361 KVSADVKKGDTPLHVAIRGRSRRRLAELLLRNPKOGRLLYRPNKAGETPNIDCSHQSKSL 420
 Qy 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKSKSLKKL 480
 Db 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKSKSLKKL 480
 Qy 481 EDEMKTFAGQOTPELPFQSWLIIVFTLLCCGGLGVPAFPVDTNLAIAISISFLALIYIF 540
 Db 481 EDEMKTFAGQOTPELPFQSWLIIVFTLLCCGGLGVFAFTVDTNLAIAISISFLALIYIF 540
 Qy 541 FIVYFGRRREGESGNNAWALSTRLARHIGYLELLFKLMFVNPPPEQTTKALPVRLF 600
 Db 541 FIVYFGRRREGESGNNAWALSTRLARHIGYLELLFKLMFVNPPPELAEQTTKALPVRLF 600
 Qy 601 TDYNRLLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEESQKKKWKTCCLPS 660
 Db 601 TDYNRLLSSVGGETSLAEMIATLSDACEREFGLATRLFRVLRTEESQKKKWKTCCLPS 660
 Qy 661 FVIFELFVGCIIAGITTLAIIFRVPDKHLTVNAILISIASVVGLAFVLCRTWQVDSLL 720
 Db 661 FVIFELFVGCIIAGITTLAIIFRVPDKHLTVNAILISIASVVGLAFVLCRTWQVDSLL 720
 Qy 721 NSQKRLHSAASKLHKLKSEGFMKLVKCEVELMARMKTIIDSFTQNTQLVVIIDGLDAC 780
 Db 721 NSQKRLHSAASKLHKLKSEGFMKLVKCEVELMARMKTIIDSFTQNTQLVVIIDGLDAC 780
 Qy 781 EQDKVLQMLDTRVRLFSKGPPIAIPASDPHIIIAINQNLNSVLDRDSNINGHDYMRNVH 840
 Db 781 EQDKVLQMLDTRVRLFSKGPPIAIPASDPHIIIAINQNLNSVLDRDSNINGHDYMRNVH 840
 Qy 841 LPVFNLSRGLSNARKFLVTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLGSKTALN 900
 Db 841 LPVFNLSRGLSNARKFLVTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLGSKTALN 900
 Qy 901 RRDYTRRRQMORTITRQMSFDLTLLVTEDFWSDISQPTMERLLNIVSVTGRLLRANQIT 960
 Db 901 RRDYTRRRQMORTITRQMSFDLTLLVTEDFWSDISQPTMERLLNIVSVTGRLLRANQIT 960
 Qy 961 FNWDRLASWINLTQWPFYRTSWLIYLEETEGLPDQMTLKIYERISKNIPTTKDVEPLL 1020
 Db 961 FNWDRLASWINLTQWPFYRTSWLIYLEETEGLPDQMTLKIYERISKNIPTTKDVEPLL 1020
 Qy 1021 BIDGDIRNFVFLSSRTPVLVARDVKTFPCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
 Db 1021 BIDGDIRNFVFLSSRTPVLVARDVKTFPCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
 Qy 1081 PPLFLHEGPPRPPSGYSQPASVCSASFGNGFPFGVSPQPHSSYSGSLGSPQHPFYNR- 1139
 Db 1081 PPLFLHEGPPRPPSGYSQPASVCSASFGNGFPFGVSPQPHSSYSGSLGSPQHPFYNR 1140
 Qy 1140 -----AAVP-----ATGSLLLSMTVD 1157
 Db 1141 FFAPLYTPRYPGGSQHILSRSSVKASLPDQNNGLPCDSGFNKRQGRSPCHRQLTVT 1200
 Qy 1158 WV-----CEKLRQIEGLDQNMPOYCTTIKKANINGRVLSCQNIDELKEMAMNFGDW 1211
 Db 1201 EFNDRGRCCKLRQIEGLDQNMPOYCTTIKKANINGRVLSCQNIDELKEMAMNFGDW 1260
 Qy 1212 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTTELPLTESSQTPYTLN 1271
 Db 1261 LFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTTELPLTESSQTPYTLN 1320
 Qy 1272 PSPELNTGLDEGAPRHNSLWSQOTRRTPSLSSLSQDSSIIISKUTDKVQAEYRDAY 1331
 Db 1321 PSPELNTGLDEGAPRHNSLWSQOTRRTPSLSSLSQDSSIIISKUTDKVQAEYRDAY 1380

QY 1332 REXIAQMSOLEGGTSGSTISGRSPHSTYYIGOSSGSGSIHSTLEQERKGEGLKQEDGR 1391
 Db 1381 REYIAQMSOLEGGTSGSTISGRSPHSTYYIGOSSGSGSIHSTLEQERKGEGLKQEDGR 1440
 QY 1392 KSFLMKRGDVIDYSSSGSVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSEPSLFTDGL 1451
 Db 1441 KSFLMKRGDVIDYSSSGSVSTNEASPLDPITEEDEKSDQSGSKLLPGKKSSEPSLFTDGL 1500
 QY 1452 KLKGGGLRYQKLPSDEDESGTGRVQIIPHCCKMIRTKLKAK-ORECASPOEHSAPERT 1510
 Db 1501 KLKGGGLRYQKLPSDEDESGTGRVQIIPHCCKMIRTKLKAK-ORECASPOEHSAPERT 1558
 QY 1511 FIKAKYLSLALLDKDSSGVSRSNESPNSHSLHNEAADDSOLEKANLIELEDEGHSGK 1570
 Db 1559 FIKAKYLSLALLDKDSSGVSRSNESPNSHSLHNEAADDSOLEKANLIELEDEGHSGK 1618
 QY 1571 RGMPHSLSGIQLDPIIARMSICSEDKSPSCSLIASPESWPACQKAYNLNRTPTVTIL 1630
 Db 1619 RGMPHSLSGIQLDPIIARMSICSEDKSPSCSLIASPESWPACQKAYNLNRTPTVTIL 1678
 QY 1631 NNNTAPNRRANQNPDEIEGIRETSQVILRPGSPNPPTAVQENILKSMHAKRSQSSYTRL 1690
 Db 1679 NNNTAPNRRANQNPDEIEGIRETSQVILRPGSPNPPTAVQENILKSMHAKRSQSSYTRL 1738
 QY 1691 SKDASELHAASSESTGFGEERESIL 1715
 Db 1739 SKDASELHAASSESTGFGEERESIL 1763

RESULT 6
 AAM39025
 ID AAM39025 standard; protein; 1715 AA.
 AC AAM39025;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2170.
 XX
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 22-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 FI WPI; 2001-442253/47.
 DR N-PSDB; AAI58181.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 4; SEQ ID NO 2170; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 1715 AA;
 Query Match 93.4%; Score 8266; DB 4; Length 1715;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
 QY 1 MSVLISQSVINYVERENIPALKALLEKCKDVBERNECGQTPLMAAEQGNVEIVKELLKN 60
 Db 1 MSVLISQSVINYVERENIPALKALLEKCKDVBERNECGQTPLMAAEQGNVEIVKELLKN 60
 QY 61 GANCNLEDLNNWTALISASKEGHIHIVEELIKSGASIEHRDMGGWTALWACVKGRTDVV 120
 Db 61 GANCNLEDLNNWTALISASKEGHIHIVEELIKSGVNLNHRDMGGWTALWACVKGRTDVV 120
 QY 121 ELLLSHGANGPSVTGLQSVYPIIWAAGRGHADIVHLLQNGAKVNGSDKYTTPLVWAAR 180
 Db 121 ELLLSHGANGPSVTGLQSVYPIIWAAGRGHADIVHLLQNGAKVNGSDKYTTPLVWAAR 180
 QY 181 KGHLECVKHLANGADVDEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
 Db 181 KGHLECVKHLANGADVDEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240
 QY 241 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 300
 Db 241 LMTASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 300
 QY 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Db 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPPYINICSHQKSIL 420
 Db 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPPYINICSHQKSIL 420
 QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLVAQWGSFLLKKL 480
 Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLVAQWGSFLLKKL 480
 QY 481 EDEMTKFAGQOETPLFQFQSWLIVFLTLCCGGGLGVFAFPVDNTLAIASLSFALIYIF 540
 Db 481 EDEMTKFAGQOETPLFQFQSWLIVFLTLCCGGGLGVFAFPVDNTLAIASLSFALIYIF 540
 QY 541 FIVYFGRRREGESWNAWALSTRLARHIGYLELLFKLMFVNPELPQOTKALPVRFLF 600
 Db 541 FIVYFGRRREGESWNAWALSTRLARHIGYLELLFKLMFVNPELPQOTKALPVRFLF 600
 QY 601 TDNRLSSVGGETSIAEMIATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCPLS 660
 Db 601 TDNRLSSVGGETSIAEMIATLSDACERBFGFLATRLFRVFRTEESQGGKKWKTCCPLS 660
 QY 661 FVIFLFIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGFLAVLNCRNTWQVLDLSL 720

Db	661	VFVFLFICGIIISGTTLLAIFRVDPKHLTVNAVLISIASVVGAFVLCNRTWQVLDLSL	720
Qy	721	NSQRKRLHSAKHLKSEGFMKVLKCEVELMARMAKTIDSFTQONQRLVVIIDGLDAC	780
Db	721	NSQRKRLHNAASKHLKSEGFMKVLKCEVELMARMAKTIDSFTQONQRLVVIIDGLDAC	780
Qy	781	EQDKVQLMDTVRVLFSGPPIAIPASDPHIIKAINQNLSVLRDSDNINGHDYMRNIVH	840
Db	781	EQDKVQLMDTVRVLFSGPPIAIPASDPHIIKAINQNLSVLRDSDNINGHDYMRNIVH	840
Qy	841	LPVFLNSGLSNARKFLYTSATNGDITCSDTTCQEDTDRRVQNSLCEMTKLGSKTALN	900
Db	841	LPVFLNSGLSNARKFLYTSATNGDITCSDTTCQEDTDRRVQNSLCEMTKLGSKTALN	900
Qy	901	RRDYRRRQMQRTITROMSFDLTKLLVTEDFWSDISPOTRMRLLNIVSVTGRLLRANOIT	960
Db	901	RRDYRRRQMQRTITROMSFDLTKLLVTEDFWSDISPOTRMRLLNIVSVTGRLLRANOIT	960
Qy	961	FNWDLASWINLTCQWYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL	1020
Db	961	FNWDLASWINLTCQWYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL	1020
Qy	1021	EIDGDIRNFVFLSRTVPLVARDVKTFPLCTVNLDPKRLREIIADVRAAREQINIGGLAY	1080
Db	1021	EIDGDIRNFVFLSRTVPLVARDVKTFPLCTVNLDPKRLREIIADVRAAREQINIGGLAY	1080
Qy	1081	PPLPLHEGPPPPSGYSPASVCSASFNPGGVSPQSHSVYSGLSGPPFPYVRA	1140
Db	1081	PPLPLHEGPPPPSGYSPASVCSASFNPGGVSPQSHSVYSGLSGPPFPYVRA	1140
Qy	1141	AVPATGSSLLSSMTVDVCEKRIQIEGLDQMMPOYCTTIKKANINGRVLSCQNIIDLK	1200
Db	1141	SGPAPGVVLNSLNDVACEKRIQIEGLDQMMPOYCTTIKKANINGRVLSCQNIIDLK	1200
Qy	1201	KEMANFGDWHLFRSMVLEMSVSQVVPDPRFLNENSSAPVPHGESARRSHTPLPT	1260
Db	1201	KEMANFGDWHLFRSTVLEMEASHVVPDPRFLSESSSGPAPGEPARRASHNELPHT	1260
Qy	1261	ELSSQTPYTLNFSFEELNLGLDEGAPRHSNLWSQOTRTPPSLSLNSQDSSIEISKLT	1320
Db	1261	ELSSQTPYTLNFSFEELNLGLDEGAPRHSNLWSQOTRTPPSLSLNSQDSSIEISKLT	1320
Qy	1321	DKVQAEYDAYREYIAQMSQLEGGTSGSTISGRSPHSHTYIGOSSSGSSTHSTLEOERG	1380
Db	1321	DKVQAEYDAYREYIAQMSQLEGGTSGSTISGRSPHSHTYIGOSSSGSSTHSTLEOERG	1380
Qy	1381	KEGELKQEDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS	1440
Db	1381	KDSBPKPDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS	1440
Qy	1441	SERPSLFQTDKLKGGGLRYOKLSDDESGTGRVQITPHCSKMRITKRLKAKORECASP	1500
Db	1441	SERPSLFQTDKLKGGGLRYOKLSDDESGTGRVQITPHCSKMRITKRLKAKORECASP	1500
Qy	1501	QEHSAEPIRTFIKAKYLSLALLDKQSDSDGVSRSNESPNSHSLHNEAADSDQLEKANLI	1560
Db	1500	PEHGAEPRTFIKAKYLSLALLDKQSDSDGVSRSNESPNSHSLHNEAADSDQLEKANLI	1560
Qy	1561	ELEDEHSGKRGMPHSLSGLODPIIARMSICSEKSPSECSLIASSPEESWPACQAYN	1620
Db	1560	ELEDDSHSGKRGMPHSLSGLODPIIARMSICSEKSPSECSLIASSPEENWPACQAYN	1619
Qy	1621	LNRTPTSTVTLNNAPTNRANQNFDEIRGIRETSQVILRPGPSPNPPTAVONENLKSMAHK	1680
Db	1620	LNRTPTSTVTLNNAPANRANQNFDEIRGIRETSQVILRPPSSSPNPTTIGNENLKSMTHK	1679
Qy	1681	RSQSSSYRLSKDASELH-AASSSTGFGGERESIL 1715	
Db	1680	RSQSSSYRLSKDPELHAAASSTGFGGERESIL 1715	

AAM38993	
ID	AAM38993 standard; protein; 1715 AA.
XX	
AC	AAM38993;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2138.
XX	
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US034263.
XX	
PR	23-DEC-1999; 99US-00471275.
PR	21-JAN-2000; 2000US-00488725.
PR	25-APR-2000; 2000US-00552317.
PR	20-JUN-2000; 2000US-00598042.
PR	19-JUL-2000; 2000US-00620312.
PR	03-AUG-2000; 2000US-00653450.
PR	14-SEP-2000; 2000US-00662191.
PR	19-OCT-2000; 2000US-00693036.
PR	29-NOV-2000; 2000US-00727344.
XX	(HYSE-) HYSEQ INC.
FA	
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;
XX	WPI; 2001-442253/47.
DR	N-PSDB; AAI58149.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
XX	
PS	Example 4; SEQ ID NO 2138; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification
XX	
SQ	Sequence 1715 AA;
XX	
Query Match	93.4%; Score 8266; DB 4; Length 1715;
Best Local Similarity	92.7%; Pred. No. 0;
Matches 1591; Conservative	53; Mismatches 70; Indels 2; Gaps 2;

Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAABQGNVEIVKELKN	60
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAABQGNLEIVKELKN	60

QY 61 GANCNLELDNWTALISASKEGHIHVEELIKSGASLEHRDMGWTALMWACYKGRDVV 120
DB 61 GANCNLELDNWTALISASKEGHVHVEELIKCGVNLLEHRDMGWTALMWACYKGRDVV 120
QY 121 ELLLSHGANSVTLQYQSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
DB 121 ELLLSHGANSVTLQYQSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
QY 181 KGHLECVKHLANGADVQSGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
DB 181 KGHLECVKHLANGADVQSGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
QY 241 LMTASKEGHIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
DB 241 LMTASKEGHTEIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKOGETPLIKATKRNIEVWELLDDKGA 360
DB 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKOGETPLIKATKRNIEVWELLDDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGRLLYRPNKAGETPPYINDCSHQKSL 420
DB 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGRLLYRPNKAGETPPYINDCSHQKSL 420
QY 421 TQIFGARHLSPTETDGMGLVDLYSSALADILSEPTMOPPICVGLYAOWGSGSKFLKKL 480
DB 421 TQIFGARHLSPTETDGMGLVDLYSSALADILSEPTMOPPICVGLYAOWGSGSKFLKKL 480
QY 481 EDEMKTFAGOOTPLFOFWSMLIVFLTLCCGLGLVFAFPVDVNLALAIISLSFLALIYIF 540
DB 481 EDEMKTFAGOOIPELFOFWSMLIVFLTLCCGLGLVFAFVHNLGIAVLSFLALIYIF 540
QY 541 FIVYFGRRREGESWNWAWLSRLARHIGVLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
DB 541 FIVYFGRRREGESWNWAWLSRLARHIGVLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
QY 601 TDYNRLLSVGGETSLAEIMATLSDACEREGFATLFRVFRTEESOGKKWKTKCCLPS 660
DB 601 TDYNRLLSVGGETSLAEIMATLSDACEREGFATLFRVFRTEESOGKKWKTKCCLPS 660
QY 661 FVIFLFTVGCIIAGITLAILFRVDPKHLTVNALISIASVVGAFVNLNCRWTMQVLDL 720
DB 661 FVIFLFTIIGCIISGITLAILFRVDPKHLTVNAVLSIASVVGAFVNLNCRWTMQVLDL 720
QY 721 NSQKRLHSAASKLHKLSGFGPKVLKCEVELMARMAKTIDSTFQNTQRLVVIIDGLDAC 780
DB 721 NSQKRLHNAASKLHKLSGFGPKVLKCEVELMARMAKTIDSTFQNTQRLVVIIDGLDAC 780
QY 781 EQDKVQLMDTVRVLFSKGPPIAFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 840
DB 781 EQDKVQLMDTVRVLFSKGPPIAFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
DB 841 LPVFLNSRGLSNARKFLVTSATNGDVCSDTTGTQEDADRRVSONSLGEMTKLGSKTALN 900
QY 901 RRDYRRRQMQRTITROMSFDLTKLVTEDEFSDISQPTWRRLLINIVSVTGRLLRANQIT 960
DB 901 RRDYRRRQMQRTITROMSFDLTKLVTEDEFSDISQPTWRRLLINIVSVTGRLLRANQIS 960
QY 961 FNWDRLASWNLTEQWYRVSLLLYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
DB 961 FNWDRLASWNLTEQWYRVSLLLYLEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNFVFLSSRTFVLVARDVKTFLPCTVNLDPKLEIIRIADVRAAREQINIGGLAY 1080
DB 1021 EIDGDIRNFVFLSSRTFVLVARDVKVFLPCTVNLDPKLEIIRIADVRAAREQISIGGLAY 1080
QY 1081 PPLPLHGGPRRPSGYSQPSVCSASFNGFPFGGVSPQPHSSYSGLSGPOHPFYNRA 1140
DB 1081 PPLPLHGGPRRPSGYSQPSVCSSTSFNGFPFAGGVSPQPHSSYSYSGMTGPOHPFYNRG 1140

QY 1141 AVPATGSSLLSSMTVDVVCCKLRQIEGLDQNNMPOYCTTIKANNINGRVLSCNIDELK 1200
DB 1141 SGPAQGVVLLSLNVDACEKLRQIEGLDQMLPOYCTTIKANNINGRVLACNIDELK 1200
QY 1201 KEMANFGDWHLFRSVMLEKRSVESQVVPEDPRFLNENSAPVPHGESARRSHTPLT 1260
DB 1201 KEMANFGDWHLFRSVMLEKRNAESHVVPEDPRFLSESSGPPHGPGEARRASHNELPHT 1260
QY 1261 ELSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSTRTTPSLNSODSIEISKLT 1320
DB 1261 ELSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSTRTTPSLNSODSIEISKLT 1320
QY 1321 DKVQAYRYDAYREYIAQMSQLEGTSSTISGRSSPHSTVYIQSSSGSGSIHSTLEOERG 1380
DB 1321 DKVQAYRYDAYREYIAQMSQLEGPGSTTISGRSSPHSTVYIMQSSSGSIHNLQOEGK 1380
QY 1381 KEGELKQBDGRKSFMLMKRGDVIDYSSGVSTNEASPLDPTITERDEKSDQSGSKLLPKKS 1440
DB 1381 KQSEPKPDGRKSFMLMKRGDVIDYSSGVSTNDASPLDPTITERDEKSDQSGSKLLPKKS 1440
QY 1441 SERPSLFOTDLKLKGGGLRYOKLPSDEDESGTGRVQITPHCSKWIRTKLAKORECASP 1500
DB 1441 SERSSLFOTDLKLKGGGLRYOKLPSDEDESGTEESDNTPLKDDKDKRKAEGKVERVPKS 1499
QY 1501 QEHSAPRTFTIKAKEYLSDALDKKSDSDSGVRSNESPNSHLSHNEAADDSOLEKANLI 1560
DB 1500 PHSAPRTFTIKAKEYLSDALDKKSDSDSGVRSSESPNSHLSHNEAVDDDSOLEKANLI 1559
QY 1561 ELEDEHSGKGRMHPHSLGSLQDPIIARMGICSDKSKSPSECSLIASSPEESWPACQKAYN 1620
DB 1560 ELEDDSHSGKGRGPHSLGSLQDPIIARMGICSDKSKSPSECSLIASSPENWPACQKAYN 1619
QY 1621 LNRTPTVTILNNTATNEANQNFDEIGIRETSQVILRPGPSPNPTAVQENLNKSMWAK 1680
DB 1620 LNRTPTVTILNNSAPANEANQNFDEIGIRETSQVILRPPSSPNPTTQENLNKSMTHK 1679
QY 1681 RSQRSSYTRLSDASELH-AAASESTGFGEREESIL 1715
DB 1680 RSQRSSYTRLSDKDPPELHAAASESTGFGEREESIL 1715
RESULT 8
AAE25144
ID AAE25144 standard; protein; 1715 AA.
AC AAE25144;
XX
DT 30-OCT-2002 (first entry)
XX Human ARMS protein.
DE
XX Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
KW ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
KW diagnostic; imaging; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 17..390
FT /note= "Ankyrin repeat"
FT Region 354..493
FT /note= "N-terminal region"
FT Domain 496..518
FT /note= "Transmembrane domain"
FT Domain 525..546
FT /note= "Transmembrane domain"
FT Region 573..638
FT /note= "Loop between TM2 and TM3"
FT Domain 661..680
FT /note= "Transmembrane domain"
FT Domain 688..710
FT /note= "Transmembrane domain"
FT Region 749..854

FT Region /note= "Carboxy tail"
 FT 940..1060
 FT Region /note= "Carboxy tail"
 FT 1081..1093
 FT Domain /note= "Polyproline stretch"
 FT 1151..1221
 FT Binding-site /note= "SAM domain"
 FT 1713..1715
 FT /note= "PDZ binding motif"
 XX
 PN WO200250273-A2.
 XX
 XX 27-JUN-2002.
 PD
 XX
 XX
 PF 19-DEC-2001; 2001WO-US048603.
 XX
 PR 21-DEC-2000; 2000US-0256909P.
 XX
 PA (UUNY) UNIV NEW YORK STATE.
 XX
 XX Chao MV, Kong H.
 PI
 XX
 DR WPI; 2002-508800/54.
 DR N-PSDB; AAD41036.
 XX
 PT Isolated ankyrin repeat-rich membrane spanning (ARMS) polypeptide that is
 PT a target for phosphorylation by neurotrophin and ephrin receptor tyrosine
 PT kinases, useful as a marker for growth cones.
 XX
 PS Claim 1; Page 103-110; 136pp; English.
 XX
 CC The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
 CC protein which is a down stream target of neurotrophin and ephrin receptor
 CC tyrosine kinases. ARMS DNA is useful for visualising the growth cone of
 CC neurons. ARMS protein is useful as an indicator of the biological
 CC activity of neurotrophins and ephrins, as a marker for neuronal cells
 CC which have the ability to undergo continued synaptic changes through
 CC adult life or for the presence and distribution of ARMS in such neuronal
 CC cells, or as a marker for growth cones. ARMS DNA is useful in diagnostic
 CC and imaging methods. The present sequence is human ARMS protein
 XX
 SQ Sequence 1715 AA;

Query Match 93.4%; Score 8266; DB 5; Length 1715;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

Qy 1 MSVLISQSVINVEENIPALKKLECKVDNERNECCQTPMLAAEQGNVEIVKELKN 60
 Db 1 MSVLISQSVINVEENIPALKKLECKVDNERNECCQTPMLAAEQGNLEIVKELKN 60
 Qy 61 GANCNLEDNDNWTALISASKEGHHIVBELLSKGSASLEHRDMGGWTALMWACYKGRDGV 120
 Db 61 GANCNLEDNDNWTALISASKEGHHIVBELLSKGVNLEHRDMGGWTALMWACYKGRDGV 120
 Qy 121 ELLLSHGANGPSVTGLQYSVYPIIWAAGRGHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
 Db 121 ELLLSHGANGPSVTGLQYSVYPIIWAAGRGHADIIVHLLLQNGAKVNCSDKYGTTPLVWAAR 180
 Qy 181 KGHLECVKHLAMGADVDEQANSMTALIVAVKGYTOSVKEILKRNPNVNLTKDGNNTA 240
 Db 181 KGHLECVKHLAMGADVDEQANSMTALIVAVKGYTOSVKEILKRNPNVNLTKDGNNTA 240
 Qy 241 LMIASKEGHHIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
 Db 241 LMIASKEGHHIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
 Qy 301 QDNKTALVWAEKGNATVRILOCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Db 301 QDNKTALVWAEKGNATVRILOCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
 Qy 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPDKGRLLYRPNKAGETTPYNDICSHQKSIL 420
 Db 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPDKGRLLYRPNKAGETTPYNDICSHQKSIL 420

Db 361 KVSADVKKGDTPLHVAIRGRSRRLAELLLRNPDKGRLLYRPNKAGETTPYNDICSHQKSIL 420
 Qy 421 TQIFGARHLSPETDGMGLGYDLYSSALADILSPTMQPPICVGLYAWGSGKSFLLKKL 480
 Db 421 TQIFGARHLSPETDGMGLGYDLYSSALADILSPTMQPPICVGLYAWGSGKSFLLKKL 480
 Qy 481 EDEMKTFAGQOQTEPLFQFSWMLIVFLTLCCGGLGVFAFPVDTNLAIAISISFLALIYIF 540
 Db 481 EDEMKTFAGQOQTEPLFQFSWMLIVFLTLCCGGLGVFAFPVDTNLAIAISISFLALIYIF 540
 Qy 541 FIVYFGRREGESNNWAWALSTRLARHIGYLELLFKLMFVNPPBELPQOTTKALPVRELF 600
 Db 541 FIVYFGRREGESNNWAWALSTRLARHIGYLELLFKLMFVNPPBELPQOTTKALPVRELF 600
 Qy 601 TDYNRLSSVGGETSLAEMIAIATSDACREFFGLATRLFRVFRTEESQCKKWKTCCLPS 660
 Db 601 TDYNRLSSVGGETSLAEMIAIATSDACREFFGLATRLFRVFRTEESQCKKWKTCCLPS 660
 Qy 661 FVIFLFIIVGCIIGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLNCRTWQVLDLIL 720
 Db 661 FVIFLFIIVGCIIGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLNCRTWQVLDLIL 720
 Qy 721 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMAMAKTIDSFQONQRLVVIDGLDAC 780
 Db 721 NSQKRLHSAASKLHLKSEGFMKVLKCEVELMARMAMAKTIDSFQONQRLVVIDGLDAC 780
 Qy 781 EODKVLQWLDTVRVLFPSKGPFIASFDPHPIIIKAINQNLNVLSDNSINGHDYMRNVH 840
 Db 781 EODKVLQWLDTVRVLFPSKGPFIASFDPHPIIIKAINQNLNVLSDNSINGHDYMRNVH 840
 Qy 841 LPVFLNRSGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
 Db 841 LPVFLNRSGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
 Qy 901 RRDYRRRQMQRTITROMSFDLTKLVTEDFWFSDISPQTMRLNLNIVSVTGLRLRANQIT 960
 Db 901 RRDYRRRQMQRTITROMSFDLTKLVTEDFWFSDISPQTMRLNLNIVSVTGLRLRANQIT 960
 Qy 961 FNWDLASWNLTEQWYRYSWMLLYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Db 961 FNWDLASWNLTEQWYRYSWMLLYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
 Qy 1021 EIDGDINFEVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIIVADVRAARQINIGGLAY 1080
 Db 1021 EIDGDINFEVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIIVADVRAARQINIGGLAY 1080
 Qy 1081 PPLPLHEGPPRPSPGYSQPAVSCSASFNGPFGGVSPQPHSSYVYSGLSGQPHFPYNRA 1140
 Db 1081 PPLPLHEGPPRPSPGYSQPAVSCSASFNGPFGGVSPQPHSSYVYSGLSGQPHFPYNRA 1140
 Qy 1141 AVPATGSSLLLSMTVDVCEKLRQIEGLDQNMMPQVCTTIKKANINGRVLSCQNIDELK 1200
 Db 1141 AVPATGSSLLLSMTVDVCEKLRQIEGLDQNMMPQVCTTIKKANINGRVLSCQNIDELK 1200
 Qy 1201 KEMANFGDWHLPFSWYLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSSHTEPLT 1260
 Db 1201 KEMANFGDWHLPFSWYLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSSHTEPLT 1260
 Qy 1261 ELSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRTRTPSLSSLSNSQSSSIEISKLT 1320
 Db 1261 ELSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSQTRTRTPSLSSLSNSQSSSIEISKLT 1320
 Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYIYGSSSGSGSHTSLEBQRG 1380
 Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYIYGSSSGSGSHTSLEBQRG 1380
 Qy 1381 KEGSLKQEDGRKSPMKRGDVIDYSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 Db 1381 KEGSLKQEDGRKSPMKRGDVIDYSSSGVSTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
 Qy 1441 SERPSLFTQDLKKGGLRYQKLPSPDESGTGRVQITPHCSKMIRTKRLKAKQRECAASP 1500
 Db 1441 SERPSLFTQDLKKGGLRYQKLPSPDESGTGRVQITPHCSKMIRTKRLKAKQRECAASP 1500
 Qy 1499 SERSSSLFTQDLKKGGLRYQKLPSPDESGTGRVQITPHCSKMIRTKRLKAKQRECAASP 1499
 Db 1499 SERSSSLFTQDLKKGGLRYQKLPSPDESGTGRVQITPHCSKMIRTKRLKAKQRECAASP 1499

QY 1501 QEHSABPRTFIRAKEVLSALLDKXSDSPGVRSSNESHSLHNEAADDSOLEKANLI 1560
DB 1500 PEHSABPRTFIRAKEVLSALLDKXSDSPGVRSSNESHSLHNEAADDSOLEKANLI 1559
QY 1561 ELDEGHSGRGRPHSLGSLQDPIIARMSICSEDKKSPSCSLIASSPEESWPACOKAYN 1620
DB 1560 ELEDDSHSGRGRPHSLGSLQDPIIARMSICSEDKKSPSCSLIASSPEENWPACOKAYN 1619
QY 1621 LNRTPTVTNNNTAPNRRANQNFDEIGRETQSQVILRPGSPNPTAVONENLKSMAHK 1680
DB 1620 LNRTPTVTNNNSAPNRRANQNFDEMEGIRETSQVILRPSPPNPTTIQENLKSMTWK 1679
QY 1681 RSQRSSVTRLSKDASELH-AASSESTGFGERSIL 1715
DB 1680 RSQRSSVTRLSKOPPELHAAASSESTGFGERSIL 1715

RESULT 9

AAU96840
ID AAU96840 standard; protein; 1715 AA.

AC AAU96840;
DT 30-JUL-2002 (first entry)
DE Human kidins220 protein.

XX Kidins220; kinase D interacting substrate of 22kDa; cytostatic;
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer;
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer;
XX human.

OS Homo sapiens.

XX WO200220786-A2.

XX 14-MAR-2002.

XX 06-SEP-2001; 2001WO-GB003977.

XX 06-SEP-2000; 2000US-0230449P.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

XX Schiavo G, Iglesias T;

XX WPI; 2002-371879/40.

XX N-PSDB; ABK51221.

XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,
PT useful for identifying modulators useful in treating cancer and
PT neurodegenerative diseases.

XX Claim 33; Fig 11; 228pp; English.

XX The invention relates to polypeptide comprising a kinase D interacting
CC substrate of 220 kDa (Kidins220) from rat and human or their fragment,
CC variant or fusion provided that the protein is not the polypeptide
CC encoded by GenBank Accession No. AB033076. Also included are a
CC polynucleotide encoding Kidins220 provided that polynucleotide is not any
CC one of the clones corresponding to the 61 GenBank Accession Nos. given in
CC the specification; an expression/replicable vector comprising the
CC polynucleotide; a recombinant host cell containing the polynucleotide or
CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220
CC ; an agent (Al) which modulates activity of Kidins220, protein kinase D
CC (PKD) or interaction between PKD and Kidins220; a mutant animal
CC transgenic for kidins220; the use of an agent capable of detecting the
CC expression of kidins220 gene in manufacture of a diagnostic reagent for
CC diagnosing or prognosing cancer or for monitoring the progression of
CC cancer in a patient; a compound comprising a moiety that selectively
CC binds to Kidins220 or its variant and another moiety; and a nucleic acid
CC molecule encoding the compound. The antibody is useful for modulating the

CC activity of Kidins220, where the antibody prevents a region of kidins220
CC interacting from another protein, or reduces the ability of kidins220 to
CC bind to adenosine triphosphate (ATP). The kidins220 polynucleotide is
CC useful for identifying an agent which modulates the promoter activity of
CC the polynucleotide, and Kidins220 is useful for identifying an agent
CC which modulates activity of Kidins220, protein kinase D (PKD) and the
CC interaction between PKD with Kidins220. Kidins220 or its polynucleotide
CC is useful in medicine, e.g. using gene therapy or for diagnosing or
CC monitoring progression of cancer in a patient. The agent is useful in
CC manufacture of medication, for use in treatment of neurodegenerative
CC disease, and the agent which inhibits function of Kidins220 gene or its
CC product is useful for treating cancer which is glioblastoma multiforme or
CC prostate cancer. The compound is useful for imaging cancer in an
CC individual, and for diagnosing or prognosing, and also for treating an
CC individual. The present sequence represents Human kidins220

XX Sequence 1715 AA;

QY Query Match 93.4%; Score 8266; DB 5; Length 1715;

DB Best Local Similarity 92.7%; Pred. No. 0;

Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

QY 1 MSVLISQSVINYVEENIPALKALLEKCKXVDNERNECGQTPLMLAAEQGNVEIVKELLKN 60

DB 1 MSVLISQSVINYVEENIPALKALLEKCKXVDNERNECGQTPLMLAAEQGNLEIVKELIKN 60

QY 61 GANCNLEDLDNWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGTDDV 120

DB 61 GANCNLEDLDNWTALISASKEGHIHVEELKSGVNLHRDMGGWTALMWACYKGTDDV 120

QY 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRCHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180

DB 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRCHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180

QY 181 KGHLECVKHLAMGADVDDQEGANSMTALIVAKGGYTQSVKILKNPNVNLTKDGN 240

DB 181 KGHLECVKHLAMGADVDDQEGANSMTALIVAKGGYTQSVKILKNPNVNLTKDGN 240

QY 241 LMTASKEGHIIVODLLDAGTYVNPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

DB 241 LMTASKEGHIIVODLLDAGTYVNPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300

QY 301 QDNKTALYWAVEKGNATWVRDIIQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360

DB 301 QDNKTALYWAVEKGNATWVRDIIQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360

QY 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGLLYRPNKAGETPNIDCSHOKSIL 420

DB 361 KVSADVKKGDTPLHVAIRGSRRLAEILLRNPKDGLLYRPNKAGETPNIDCSHOKSIL 420

QY 421 TQIFGARHLSPTETDGMGLDYLLSSALADIISEPTMQPPICVGLYAQSGSGSFLLKLL 480

DB 421 TQIFGARHLSPTETDGMGLDYLLSSALADIISEPTMQPPICVGLYAQSGSGSFLLKLL 480

QY 481 EDEMKTFAQOOTEPLFQFSLNIVFLTLCCGGLGVFAPVDTNLAIAISLSLALYIF 540

DB 481 EDEMKTFAQOOTEPLFQFSLNIVFLTLCCGGLGVFAPVDTNLAIAISLSLALYIF 540

QY 541 FIVIVFGRREGESWNWAWALSTRLARHIGYLELLPKLMFVNPPPEOTTKALPVRF 600

DB 541 FIVIVFGRREGESWNWAWALSTRLARHIGYLELLPKLMFVNPPPEOTTKALPVRF 600

QY 601 TDYNRLLSSVGGTSLAEMTATISDACERFGLATFLFRVFEETESQKKWKTKCCLPS 660

DB 601 TDYNRLLSSVGGTSLAEMTATISDACERFGLATFLFRVFEETESQKKWKTKCCLPS 660

QY 661 FVIFLFIIVGCIAGITLLAI FRVDPKHLTVNAILISIASVVGGLAFVLCNRTWQVLDL 720

DB 661 FVIFLFIIVGCIAGITLLAI FRVDPKHLTVNAILISIASVVGGLAFVLCNRTWQVLDL 720

QY 721 NSQRKELHSAASKLHLKSGEGFMKVLKCEVELMARMAKTIDSTQNTQRLVVIIDGLDAC 780

DB 721 NSQRKELHNAASKLHLKSGEGFMKVLKCEVELMARMAKTIDSTQNTQRLVVIIDGLDAC 780

QY 781 EODKVLQMLDTRVRLFSKGPFTAFASDPHIIIIKAINQNLSVLRSDNSINGHDYMRNIWH 840
Db 781 EODKVLQMLDTRVRLFSKGPFTAFASDPHIIIIKAINQNLSVLRSDNSINGHDYMRNIWH 840
QY 841 LPVFLNSGLSNARKFLVTSATNGDITCSDDTGTQEDTDRVSQNSLGMTKLSKLTALN 900
Db 841 LPVFLNSGLSNARKFLVTSATNGDITCSDDTGTQEDTDRVSQNSLGMTKLSKLTALN 900
QY 901 RRDYRRRQMQRTTITROMSFOLTKLLVTFEDFSDISPTMRRLNIVSVTKLLRANQIT 960
Db 901 RRDYRRRQMQRTTITROMSFOLTKLLVTFEDFSDISPTMRRLNIVSVTKLLRANQIS 960
QY 961 FNVDRLASWINLQWPTYSWMLLYLEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 FNVDRLASWINLQWPTYSWMLLYLEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNFVFLSRTPVLVARDVKFPLCTVNLDPKRLRIIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSRTPVLVARDVKFPLCTVNLDPKRLRIIADVRAAREQISIGGLAY 1080
QY 1081 PPLPLHEGPPPPSGYSPASVCSASFNGPFGVSPQPHSSVYSGLSGPQHPFYNRA 1140
Db 1081 PPLPLHEGPPPPSGYSPASVCSASFNGPFGVSPQPHSSVYSGLSGPQHPFYNRG 1140
QY 1141 AVPATGSLLSMTVDVVCCKLRQIEGLDQMMPOYCTTIKKANINGRVLSCQNIDELK 1200
Db 1141 SGAPGPPVLLNSLNDVACEKLRQIEGLDQMLPOYCTTIKKANINGRVLACQNIDELK 1200
QY 1201 KEMANFGDMLFRSVMLEMSVSQVVPFRLNENSSAPVPHGESARRSHTPLT 1260
Db 1201 KEMANFGDMLFRSVMLEMSVSQVVPFRLNENSSAPVPHGESARRSHTPLT 1260
QY 1261 ELSSQTPYTLNFSFEELNLGLDGAPRHSNLWSQSTRTPPSLSLNSQDSSIEISKLT 1320
Db 1261 ELSSQTPYTLNFSFEELNLGLDGAPRHSNLWSQSTRTPPSLSLNSQDSSIEISKLT 1320
QY 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSSGSIHSTLEQERG 1380
Db 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSSGSIHSTLEQERG 1380
QY 1381 KEGELKEDGRKSFMLKRGDVIDYSSGVSNTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
Db 1381 KDSBPKPDDGRKSFMLKRGDVIDYSSGVSNTNEASPLDPTTEDEKSDQSGSKLLPGKKS 1440
QY 1441 SERPSLFTDLKLGGLRYOKLPSEDESGTGRVQITPHCSKMYRTKLLKAKQRECAPS 1500
Db 1441 SERPSLFTDLKLGGLRYOKLPSEDESGTGRVQITPHCSKMYRTKLLKAKQRECAPS 1500
QY 1501 QEHSAEPIRTFIKAEYLSDALDKDSSDSGVRSNNESSPNHSLHNEAADDQLEKANLI 1560
Db 1500 PEHSAEPIRTFIKAEYLSDALDKDSSDSGVRSNNESSPNHSLHNEAADDQLEKANLI 1559
QY 1561 ELEDEHSGKRGPHSLSGLODPIIARMSICSEDKGSPSECSLIASSPESWPACQAYN 1620
Db 1560 ELEDDSHSGKRGPHSLSGLODPIIARMSICSEDKGSPSECSLIASSPESWPACQAYN 1619
QY 1621 LNRTPSTVTLNNAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENIKSMAHK 1680
Db 1620 LNRTPSTVTLNNAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENIKSMAHK 1679
QY 1681 RSQRSSVTRLSKDASELH-AASSESTGFGERESIL 1715
Db 1680 RSQRSSVTRLSKDPPELHAASSESTGFGERESIL 1715

RESULT 10
AAE32128
ID AAE32128 standard; protein; 1715 AA.
XX
AC AAE32128;
XX
DT 24-MAR-2003 (first entry)

XX Human cytoskeleton-associated protein, CSAP-26.
DE Human; cytoskeleton-associated protein; CSAP-26; atherosclerosis; cancer;
KW gene therapy.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Domain 42..96
FT Domain /note= "Domain present in ZO-1"
FT Domain 354..392
FT Domain /note= "Domain present in ZO-1"
FT Binding-site 467..474
FT Domain /note= "ATP/GTP-binding site motif A"
FT Domain 494..514
FT Domain /note= "Transmembrane domain"
FT Domain 524..544
FT Domain /note= "Transmembrane domain"
FT Domain 654..674
FT Domain /note= "Transmembrane domain"
FT Domain 687..707
FT Domain /note= "Transmembrane domain"
XX WO200279404-A2.
PN 10-OCT-2002.
XX 25-MAR-2002; 2002WO-US009288.
XX 29-MAR-2001; 2001US-0280508P.
PR 03-APR-2001; 2001US-0281323P.
PR 13-APR-2001; 2001US-0283769P.
PR 04-MAY-2001; 2001US-0288609P.
PR 10-MAY-2001; 2001US-0290518P.
PR 18-MAY-2001; 2001US-0291870P.
PR 29-MAY-2001; 2001US-0294451P.
XX (INCY-) INCYTE GENOMICS INC.
XX Hafalia AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
PI Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y;
PI Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
PI Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
PI Walia NK, Griffin JA, Lee EA, Swarnakar A, Ring HZ, Jones KA;
XX WPI; 2003-092894/08.
DR N-PSDB; AAD49615.
XX New human cytoskeleton-associated proteins, useful for preparing a
PT composition for diagnosing or treating a disease or condition associated
PT with decreased expression or overexpression of functional CSAP e.g.,
PT cancer.
XX Claim 1; Page 195-199; 233pp; English.
XX The invention relates to new human cytoskeleton-associated protein (CSAP)
CC and its polynucleotide. The polypeptide is useful for preparing a
CC composition for diagnosing or treating a disease or condition associated
CC with decreased expression or overexpression of functional CSAP e.g.
CC atherosclerosis or cancer. The present sequence is human CSAP-26 protein.
CC The invention is useful in gene therapy
XX
SQ Sequence 1715 AA;
Query Match 93.4%; Score 8266; DB 6; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
QY 1 MSVLISQSVINYVEENIPALKALLEKCDVDERNECGQTPMLAAECQNVIEIKELKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCDVDERNECGQTPMLAAECQNVIEIKELKN 60

Qy	61	GANCNLEDLNDWTALISASKEGHIHIVEELLKSGASLEHRDMGSGWTALMWACYKGRD	120
Db	61	GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHRDMGWTALMWACYKGRD	120
Qy	121	ELLLSHGANSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVAA	180
Db	121	ELLLSHGANSVTGLQYSVYPIIWAAGRGHADIVHLLLQNGAKVNSDKYGTTPLVAA	180
Qy	181	KGLECVKHLWAGADVQEGANSMTALIVAVKGGYTQSKEILKRNPNVNLTKDGN	240
Db	181	KGLECVKHLWAGADVQEGANSMTALIVAVKGGYTQSKEILKRNPNVNLTKDGN	240
Qy	241	LMIASKEGHIEIVODLLDAGTYNIPDRSGDTVLVIGAVRGHVEIVRALLQKYAD	300
Db	241	LMIASKEGHTEIVODLLDAGTYNIPDRSGDTVLVIGAVRGHVEIVRALLQKYAD	300
Qy	301	QDNKTALYWAVEKGNATWVRDIILOCPDTEICTKOGTPTLIKATKRNIEVBELLDK	360
Db	301	QDNKTALYWAVEKGNATWVRDIILOCPDTEICTKOGTPTLIKATKRNIEVBELLDK	360
Qy	361	KVSAVDKKGDTPLHVAIRGRSRRLAEILLRNPKDGLLYEPNKAGETPYNIDCSH	420
Db	361	KVSAVDKKGDTPLHVAIRGRSRRLAEILLRNPKDGLLYEPNKAGETPYNIDCSH	420
Qy	421	TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQMGSGK	480
Db	421	TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQMGSGK	480
Qy	481	EDEWKTFAQOOTPELPFOFSWLIIVFLTLLCGGLGVPAFPVDTNLAIAISLSF	540
Db	481	EDEWKTFAQOOTPELPFOFSWLIIVFLTLLCGGLGVPAFPVDTNLAIAISLSF	540
Qy	541	FIVYFGRRREGESWNWAWLSLRLARHIGVLELLFKLMFVNPPEIQOTTKALPV	600
Db	541	FIVYFGRRREGESWNWAWLSLRLARHIGVLELLFKLMFVNPPEIQOTTKALPV	600
Qy	601	TDYNRLSSVGGETSIAEMIATLSDACERBFGFLATLFRVFRTEESOGKKWKTC	660
Db	601	TDYNRLSSVGGETSIAEMIATLSDACERBFGFLATLFRVFRTEESOGKKWKTC	660
Qy	661	FVIFLFTIVGCIIGITITLAIIFRVPDKHLTVNAILISIASVVGIAFVLNCR	720
Db	661	FVIFLFTIVGCIIGITITLAIIFRVPDKHLTVNAILISIASVVGIAFVLNCR	720
Qy	721	NSQKRKLHSAASKHLKLSGEMKVKCEVELMARMAKTIDSFQNTROLVLIIDGL	780
Db	721	NSQKRKLHSAASKHLKLSGEMKVKCEVELMARMAKTIDSFQNTROLVLIIDGL	780
Qy	781	EQDKVLOMLDTVRVLFSGGPFIAIFASDPHIIKAINQNLSVLRNSNINGHDM	840
Db	781	EQDKVLOMLDTVRVLFSGGPFIAIFASDPHIIKAINQNLSVLRNSNINGHDM	840
Qy	841	LPVFLNSRGLSNARKPLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTK	900
Db	841	LPVFLNSRGLSNARKPLVTSATNGDVPSCDTTGTQEDADRVRVSONSLGEMTK	900
Qy	901	RRUTYRRRQORITITROMSPDLTKLIVTEDFWSDISQOTWRLLINVSVTGR	960
Db	901	RRUTYRRRQORITITROMSPDLTKLIVTEDFWSDISQOTWRLLINVSVTGR	960
Qy	961	FNWDRLASWNLTEOWPYRISWLIILEETEGLPDQMTLMTYERISKNIPTTKD	1020
Db	961	FNWDRLASWNLTEOWPYRISWLIILEETEGLPDQMTLMTYERISKNIPTTKD	1020
Qy	1021	EIDGDIRNFVEYFSSRTPVLVARDVKTFLPCTVNLDPKLEIIRIADVRAARE	1080
Db	1021	EIDGDIRNFVEYFSSRTPVLVARDVKTFLPCTVNLDPKLEIIRIADVRAARE	1080
Qy	1081	PPLPLHGGPRRPSGYQOPASVCSASFNGPPFGVGVSPPHSSYSYSGLSGPO	1140
Db	1081	PPLPLHGGPRRPSGYQOPASVCSSTSFNGPPFGVGVSPPHSSYSYSGMTGP	1140
Qy	1141	AVPATSGSLLLSMTVDVWCVKLROIENGLQDNMMPOYCTTIKCANINGRVL	1200

Db	1141	SGPAPGVLLNSLNDVACEKQLQIEGLDQSMLPQYCTTIKANTNGRVLAQCNTDELK	1200
Qy	1201	KEMAMFGDWHLFRSMVLEMRSVESOVPEDPRLFNENSAPVPHGESARRSSHTLPLT	1260
Db	1201	KEMNMGDWHLFRSTVLEMRNASHVVPEDPRFLSESSGPAPHGEPAKRAHSHNLPHT	1260
Qy	1261	ELSSQTPYTLNLFSEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNQDSSIEISKLT	1320
Db	1261	ELSSQTPYTLNLFSEELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSLNQDSSIEISKLT	1320
Qy	1321	DKVOAEYRDAYRYIAQMSOLEGGTGSSTISGRSSPHSTVYIQSSSGGSIHSTLQBERG	1380
Db	1321	DKVOAEYRDAYRYIAQMSOLEGGTGSSTISGRSSPHSTVYIMQSSSGGSIHSLNLEQEK	1380
Qy	1381	KEGELKQEDGRKSFMLMKRGDIVDYSYGVSSTNEASPLDPTTEDEKSDQSGSKLLPGKKS	1440
Db	1381	KDSEPKPDDGRKSFMLMKRGDIVDYSYGVSSTNDASPLDPTTEDEKSDQSGSKLLPGKKS	1440
Qy	1441	SERPISLFTQDLKLGGLGRYQKLPSEDESGTGRVQITPHCSKMWIRTKRLKAKQRECA	1500
Db	1441	SERSSLFQTDLKLKGLGRYQKLPSEDESGTGRVQITPHCSKMWIRTKRLKAKQRECA	1500
Qy	1501	QESASPIRFTFIKAEVYLSDALLDKKDSDSGVRSNESPNSHLSHNEAADDSOLEKANLI	1560
Db	1500	PEHSASPIRFTFIKAEVYLSDALLDKKDSDSGVRSNESPNSHLSHNEAADDSOLEKANLI	1559
Qy	1561	ELDEGHSGKRGMPHSLSGLQDPIIARMSTCSDEKSKSECSLIASSPESWPACOKAYN	1620
Db	1560	ELDDSHSGKRGPIPHLSLSGLQDPIIARMSTCSDEKSKSECSLIASSPESWPACOKAYN	1619
Qy	1621	LNRTPTSTVTLNNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPNTAVQENLKSMAHK	1680
Db	1620	LNRTPTSTVTLNNNSAPANRANQNFDEMEGIRETSQVILRPPSSSPNPTTIQENLKSMT	1679
Qy	1681	RSQRSSYTRLKSDASELH-AASSESTGFGEERESIL	1715
Db	1680	RSQRSSYTRLKSDPPPELHAAASSESTGFGEERESIL	1715
RESULT 11			
ADA09888			
ID	ADA09888	standard; protein; 1753 AA.	
AC	ADA09888;		
DT	20-NOV-2003	(first entry)	
DE	Human receptor and membrane-associated protein REMAP-30, SEQ ID:30.		
KW	Human; receptor and membrane-associated protein; REMAP;		
KW	cell proliferative disorder; cancer; autoimmune disorder;		
KW	inflammatory disorder; infection; neurological disorder;		
KW	metabolic disorder; developmental disorder; endocrine disorder;		
KW	cytostatic; immunosuppressive; antiinflammatory; neuroprotective;		
KW	nootropic; cerebroprotective; gene therapy; receptor.		
OS	Homo sapiens.		
FN	WO2003070902-A2.		
PD	28-AUG-2003.		
PF	18-FEB-2003; 2003WO-US004902.		
PR	20-FEB-2002; 2002US-0358279P.		
PR	13-MAR-2002; 2002US-0364338P.		
PR	25-APR-2002; 2002US-0375657P.		
PR	29-APR-2002; 2002US-0376669P.		
PR	10-MAY-2002; 2002US-0379837P.		
PR	10-MAY-2002; 2002US-0379853P.		
PA	(INCY-) INCYTE GENOMICS INC.		

XX Chawla NK, Yue H, Richardson TW, Marquis JP, Lehr-Mason PM;
PI Gorvad AE, Becha SD, Kable AE, Swarnakar A, Jin P, Hawkins PR;
PI Chien D, Ramkumar J, Tran UK, Hafalia AJA, Baughn MR, Lee SY;
PI Jiang X, Jackson AA, Khare R, Bulloch SA;
XX WPI; 2003-697610/66.
DR N-PSDB; ADA09926.
XX
PT New human receptor and membrane associated proteins and nucleic acids,
PT useful for diagnosing, treating or preventing e.g. viral, bacterial,
PT fungal, parasitic, protozoan or helminthic infections, cancers,
PT neurological disorders.
XX
PS Claim 1; Page 254-257; 298pp; English.
XX
CC The invention relates to 38 human receptors and membrane-associated
CC proteins (REMAP), REMAP-1 to REMAP-38, and the cDNAs encoding them
CC (ADA09859-ADA09934). The invention also encompasses expression
CC constructs, host cells and transgenic organisms comprising a REMAP
CC nucleic acid sequence; the recombinant preparation of a REMAP; an
CC antibody against a REMAP; methods of detection of REMAP proteins or
CC nucleic acids; a micro-array containing REMAP nucleic acids; methods of
CC screening compounds for their ability to modulate REMAP activity or
CC expression; and pharmaceutical compositions comprising a REMAP protein, a
CC REMAP agonist or REMAP antagonist. The REMAP proteins, nucleic acids or
CC compositions comprising them are useful in diagnosing, treating or
CC preventing a variety of disorders, including cell proliferative disorders
CC (e.g., arteriosclerosis, cirrhosis, hepatitis, psoriasis, or primary
CC thrombocytopenia) or cancers (e.g., adenocarcinoma, leukemia, or cancers
CC of the bone, brain, breast or uterus); autoimmune/inflammatory disorders
CC (e.g., AIDS, allergies, anaemia, asthma, bronchitis, gout, multiple
CC sclerosis, osteoarthritis, rheumatoid arthritis, or systemic lupus
CC erythematosus); viral, bacterial, fungal, parasitic, protozoan or
CC helminthic infections; neurological disorders (e.g., epilepsy, stroke,
CC Alzheimer's disease, dementia, or Parkinson's disease); metabolic
CC disorders (e.g., hereditary fructose intolerance, obesity, hypoglycaemia
CC or diabetes); developmental disorders (e.g., achondroplastic dwarfism,
CC hypothyroidism or hydrocephalus); or endocrine disorders (e.g., disorders
CC of the hypothalamus or pituitary gland). The present sequence represents
CC a REMAP of the invention.
XX
SQ Sequence 1753 AA;

Query Match 93.0%; Score 8237; DB 6; Length 1753;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 40; Gaps 3;

Qy 1. MSVLISQSVINVEENIPALKALLECKVDNERNECQTPMLAAEQGNVEIVKELIKN 60
Db 1 MSVLISQSVINVEENIPALKALLECKVDNERNECQTPMLAAEQGNVEIVKELIKN 60

Qy 61 GANCNLELDNWTALISAKSGHIIHVEELKSGASLEHRDMGGWTALMWACYGRITDVV 120
Db 61 GANCNLELDNWTALISAKSGHIIHVEELKSGVLEHRDMGGWTALMWACYGRITDVV 120

Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHLLONGAKVNCSDKYCTTPLVWAAR 180
Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHLLONGAKVNCSDKYCTTPLVWAAR 180

Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTOSVKEILKRNPNVNLTDKGNNTA 240
Db 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTOSVKEILKRNPNVNLTDKGNNTA 240

Qy 241 LMIASKEGHIIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVIRVALLQYADIDIRG 300
Db 241 LMIASKEGHEIIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVIRVALLQYADIDIRG 300

Qy 301 QDNKTALYWAEGKGNATWVRDILOCNPDTEICTKDGETPLIKATMNRNIEVVELLDKGA 360
Db 301 QDNKTALYWAEGKGNATWVRDILOCNPDTEICTKDGETPLIKATMNRNIEVVELLDKGA 360

Qy 361 KVSADVKKGGTPLHVAIRGRSRLAELLRNPKDGRLLYRNKAGETPTYNIDCSHQSIL 420

Db 361 KVSADVKKGGTPLHVAIRGRSRLAELLRNPKDGRLLYRNKAGETPTYNIDCSHQSIL 420
Qy 421 TQIFGARHLSPTETDGMGLYSSALADILSPTWQPPICVGLYAQWGSGKSFLLKKL 480
Db 421 TQIFGARHLSPTETDGMGLYSSALADILSPTWQPPICVGLYAQWGSGKSFLLKKL 480
Qy 481 EDEMKTFAGQOTPELFOFWSMLIVELTLLCCGGLGVFAFPVDTNLAIASISFLALYIF 540
Db 481 EDEMKTFAGQOIEPLFOFWSMLIVELTLLCCGGLGVFAFTVHPNLGIAVSFLALYIF 540
Qy 541 FIVYFGRRREGESNNWAWALSTRLARHIGYLELLFKLMFVNPPPELPTTKALPVRELF 600
Db 541 FIVYFGRRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPPELPTTKALPVRELF 600
Qy 601 TDYNRLSVGGETSIAEMIATLSACERERGFATRLFRVPTESQKKKKKTKCCLPS 660
Db 601 TDYNRLSVGGETSIAEMIATLSACERERGFATRLFRVFEKTEDTQKKKKKTKCCLPS 660
Qy 661 FVIFLFTVGCIAGITLLAIFRVDPKHLTVNAIISIASVVGFLAFVLCNRTWQVLDL 720
Db 661 FVIFLFIIGCIISITLLAIFRVDPKHLTVNAVLIISIASVVGFLAFVLCNRTWQVLDL 720
Qy 721 NSQKRLHSAASKLHLKSEGMKVLKCEVELMARMARKTIDSTFQNTQRLVVIIDGLDAC 780
Db 721 NSQKRLHNAASKLHLKSEGMKVLKCEVELMARMARKTIDSTFQNTQRLVVIIDGLDAC 780
Qy 781 EODKVLQMDTVRVLFSGKPPAIASFADPHIIKAIQNQLNSVLNDSNINCHDYMNRIVH 840
Db 781 EODKVLQMDTVRVLFSGKPPAIASFADPHIIKAIQNQLNSVLNDSNINCHDYMNRIVH 840
Qy 841 LPVFLNRSGLSNARKFLVTSATNGDITCSDTTGQEDTDRRVSONSLGEMTKLASKTALN 900
Db 841 LPVFLNRSGLSNARKFLVTSATNGDVPDCTDTTGQEDADRRVSONSLGEMTKLASKTALN 900
Qy 901 RRDYTRRRQMQRTITRQMSFDLTLLVTEDFWSDISPTQWRLNINVSVTGRLLRANOIT 960
Db 901 RRDYTRRRQMQRTITRQMSFDLTLLVTEDFWSDISPTQWRLNINVSVTGRLLRANQIS 960
Qy 961 FNWDLASWINLTQWPFYRTSWLILYEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLTQWPFYRTSWLILYEETEGIPDQMTLKTMYERISKNIPTTKDVEPLL 1020

Qy 1021 EIDGDIENFEVFLSSRTPVLVARDVKTFLCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIENFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIIADVRAAREQISIGGLAY 1080

Qy 1081 PPLPLHEGPPRPSPGYSQASVCSASFNGPFPQGVVSPQPHSSYYSGLSQPHFPYNR - 1139
Db 1081 PPLPLHEGPPRPSPGYSQSPVCSSTSFNGPFPQGVVSPQPHSSYYSGMTGPHFPYNR 1140

Qy 1140 -----AAMPATGSSLLLSMTVDVUCEK 1162
Db 1141 FFAPLYTPRYPGSQHLLSRPSVKTSPLRQDNQNGLSGSPAGPVVVLNLSLNDVACEK 1200

Qy 1163 LRQIEGLDQNMPOYCTTIKKANINGRVLSCNIDELKEMANMFGDWHLFRSVVLEMR 1222
Db 1201 LKQIEGLDQNMPOYCTTIKKANINGRVLACNIDELKEMNMFQDWHLFRSTVLEMRN 1260

Qy 1223 VESOVVPEDPRLFNENSSAPVPHGESARRSHTLPLTELSSQTPYTLNFSFEELNTLGL 1282
Db 1261 AESHVVPEDPRLFESSSSGAPGEPARRASHNELPHTLSSQTPYTLNFSFEELNTLGL 1320

Qy 1283 DEGAHRSNLSWSQOTRTPSLSLNSQDSSEIISKUTDKVQABYRAYREYIAQMSOLE 1342
Db 1321 DEGAHRSNLSWSQOTRTPSLSLNSQDSSEIISKUTDKVQABYRAYREYIAQMSOLE 1380

Qy 1343 GGTGSSITISGRSSPHSTYTYGQSSSGGSIHSTLEOERKSGELKQEDGRKSFLMKRGDI 1402
Db 1381 GGTGSSITISGRSSPHSTYTYMGQSSSGGSIHSTLEOERKSGELKQEDGRKSFLMKRGDI 1440

Qy 1403 DYSSSGVSWTNEASPLDPTBEDEKSDQSGSKLAPGKKSRRPSLFTQDLKLGCGGLRYQK 1462

Db 1441 DYSSGVTNDASPLDPIEDKSDQSGSKLPGKKSERSLSFQTDLLKXGSLRYOK 1500
Qy 1463 LPSEDESGTRVOITPHCSKMIRTKRLKAKQREKASPOHSAEPITRTFKAKYELSDAL 1522
Db 1501 LPSEDESGTEESDNTS-LLKDDKDKAEGKVERVPKSPHSAEPITRTFKAKYELSDAL 1559
Qy 1523 LDKKSDSGVRNESPNSHLNEAADDQLEKANIIELEDHSGKRGMPHSLGLOD 1582
Db 1560 LDKKSDSGVRNESPNSHLNEAADDQLEKANIIELEDHSGKRGIPHSLSGLOD 1619
Qy 1583 PIARMSICSEDKSPSECSLIASSPESWPAQKAYNLNRTPTSTVTLNNTAPTNRANQ 1642
Db 1620 PIARMSICSEDKSPSECSLIASSPESWPAQKAYNLNRTPTSTVTLNNSAPANRANQ 1679
Qy 1643 NFDEIGIRTSQVILRPGSPNPTAVONENLKSMAHRSQRSYTRLSKADSELH-AAS 1701
Db 1680 NFDEMEGIRETSQVILRPPSSPNPTTQENELKSMTHKRSQRSYTRLSKDPPELHAAS 1739
Qy 1702 SESTGFEERESIL 1715
Db 1740 SESTGFEERESIL 1753

RESULT 12
ADC06847
ID ADC06847 standard; protein; 1771 AA.
XX AC
AC ADC06847;
DT 18-DEC-2003 (first entry)
XX AC
XX Human Kidins220Pc protein AB033076.
XX
XX cytosolic; prostate cancer; breast; gene therapy; transgenic; human;
KW Kidins220Pc; kinase D-interacting substrate of 220kDa; chromosome 2p25.1.
XX
XX Homo sapiens.
XX W02003064599-A2.
XX
XX 07-AUG-2003.
XX
XX 24-JAN-2003; 2003WO-US001943.
XX
XX 25-JAN-2002; 2002US-00054935.
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00141198.
PR 19-JUL-2002; 2002US-00197824.
XX
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX WPI; 2003-679495/64.
XX
XX New isolated polynucleotide related to cancer genes, useful for
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
PT or treating cancers, e.g. breast and prostate cancers.
XX
XX Disclosure; Fig 20; 128pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a
XX differentially-regulated mammalian cancer gene. The polynucleotides of
XX the invention demonstrate cytostatic activity and are differentially
XX expressed in prostate cancer. The polynucleotide, polypeptides and
XX methods of the invention may be useful for detecting, diagnosing,
XX staging, monitoring, prognosticating, preventing or treating cancers,
XX particularly breast and prostate cancers. Furthermore, the invention may
XX be utilised during gene therapy procedures or in the production of
XX transgenic animals. The current sequence is that of the prostate cancer-
XX related protein of the invention. The current sequence is not fully

CC legible within the specification; the complete sequence was obtained from
CC GenBank.
XX
SQ Sequence 1771 AA;
Query Match 92.8%; Score 8212; DB 7; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;
Qy 1 MSVLSQSIVNVYVEENIPALKALLEKCKDVDERNECGQTPMLAAEOGNEVIVKELLKN 60
Db 1 MSVLSQSIVNVYVEENIPALKALLEKCKDVDERNECGQTPMLAAEOGNEVIVKELLKN 60
Qy 61 GANCNLEDDNWTALTALSASKEGHIHIVEILLKSGASLEHRDMGWTALMAWACYKGRDVV 120
Db 61 GANCNLEDDNWTALTALSASKEGHVHIVEILLKCGVNLHRDMGWTALMAWACYKGRDVV 120
Qy 121 ELLLSHGANSPTGLQYVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVMAAR 180
Db 121 ELLLSHGANSPTGL-YSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVMAAR 179
Qy 181 KGLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKRNPNVNLTKDGNTA 240
Db 180 KGLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILKRNPNVNLTKDGNTA 239
Qy 241 LMTASKEGHIIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 300
Db 240 LMTASKEGHEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 299
Qy 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPLHVAIRGRRLABELLNRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHVAIRGRSRKLAELLNRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 419
Qy 421 TQIFGARHLSPTETDGMGLDYSSALADILSEPTWQPPICVGLYAQWGSFGSKFLKKL 480
Db 420 TQIFGARHLSPTETDGMGLDYSSALADILSEPTWQPPICVGLYAQWGSFGSKFLKKL 479
Qy 481 EDEMTKTFAGQTEPLFQFQSWLIVFLTLCCGGGLGVFAFPVDTNLAIAISLSFLALIYIF 540
Db 480 EDEMTKTFAGQTEPLFQFQSWLIVFLTLCCGGGLGVFAFPVDTNLAIAISLSFLALIYIF 539
Qy 541 FIVIFYGRRREGESWNWALSTRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRF 600
Db 540 FIVIFYGRRREGESWNWALSTRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRF 599
Qy 601 TDYNRLSSVGGTSLAEMIATLSACERBFGFLATRLFRVFRTEESQKKKKWKTCCCLPS 660
Db 600 TDYNRLSSVGGTSLAEMIATLSACERBFGFLATRLFRVFRTEESQKKKKWKTCCCLPS 659
Qy 661 FVIFLEFIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGFLAVNLCTRWQVLDL 720
Db 660 FVIFLEFIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGFLAVNLCTRWQVLDL 719
Qy 721 NSQKRLHSAASKLHKLKSEGFNKLKCEVELMARMAKTIDSTONOTRLVLIIDGLDAC 780
Db 720 NSQKRLHSAASKLHKLKSEGFNKLKCEVELMARMAKTIDSTONOTRLVLIIDGLDAC 779
Qy 781 EQDKVLQMLDTRVRLFSKGPFIIFASDPHIIKAINONLNSVLRDSNINNGHYMRNVH 840
Db 780 EQDKVLQMLDTRVRLFSKGPFIIFASDPHIIKAINONLNSVLRDSNINNGHYMRNVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSTPTTQEDTDRVSONSLGEMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSTPTTQEDTDRVSONSLGEMTKLGSKTALN 899
Qy 901 RRDYRRRQMRITRQMSFDLTKLIVTEWFSDI SPQTRRLNIVSVTGRLLRANQIT 960
Db 900 RRDYRRRQMRITRQMSFDLTKLIVTEWFSDI SPQTRRLNIVSVTGRLLRANQIS 959

QY 961 FNDRLASWINTQWYRTSMILLYLEBTEGLPQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 960 FNDRLASWINTQWYRTSMILLYLEBTEGLPQMTLKTMYERISKNIPTTKDVEPLL 1019
QY 1021 EIDGDIRPEVLSRTPLVARDVKTFPLCTVNDLPKLEIADVRAAREQINTGGLAY 1080
Db 1020 EIDGDIRPEVLSRTPLVARDVKTFPLCTVNDLPKLEIADVRAAREQINTGGLAY 1079
QY 1081 PPLPLHEGPPRPSPGYSOPASVCSASFNGPPGVSPQPHSSVYSGLSGPHQPFYN-- 1138
Db 1080 PPLPLHEGPPRPSPGYSOPASVCSASFNGPPGVSPQPHSSVYSGLSGPHQPFYNRP 1139
QY 1139 -----RAAVP 1143
Db 1140 FPAFLYTPRYYPGSOHLISRPVSKTSLPRDQNNGLEVIKEDAAEGLSSPTDSRGSGP 1199
QY 1144 ATGSSILLSSMTVDVCEKLOIEGLDQNMPOYCTTIKKANINGRVLSCNIDELKKEM 1203
Db 1200 APGVVLLNSLVDAVCEKLOIEGLDQMLPOYCTTIKKANINGRVLACQCNIDELKKEM 1259
QY 1204 AMNFGDMLFRSMVLEMRSVESQVPPEDPRFLNENSSAPVPHGESARRSSHTELPLELS 1263
Db 1260 NMNFGDMLFRSTVLEMRNAESHVVPEDPRFLSESSEGPAPHGEPARRASHNELPHELS 1319
QY 1264 SQTPTLNFSPFELNLGLDEGAPRHSNLWSQOTRTPSLSSLSQDSSIEISKLTDKV 1323
Db 1320 SQTPTLNFSPFELNLGLDEGAPRHSNLWSQOTRTPSLSSLSQDSSIEISKLTDKV 1379
QY 1324 QAEYRDAYREYIAQMSQLEGGTSGTISGRSPHSTYIYGSSSGGSHSTLEQBRGEG 1383
Db 1380 QAEYRDAYREYIAQMSQLEGGTSGTISGRSPHSTYIYGSSSGGSHSTLEQBRGKDS 1439
QY 1384 ELKQDGRKSLFKRKGVDIVSSGVSTNEASPLDPITEEDEKSDQSGSKLLPGKSSER 1443
Db 1440 EPKPDGRKSLFKRKGVDIVSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKSSER 1499
QY 1444 PSLFTDMLKGGGLRYOKLPSDEDESGTGRVQIIPHCSKMIIRTKLKAQREKASPOEH 1503
Db 1500 SSLFTDMLKGGGLRYOKLPSDEDESGTGRVQIIPHCSKMIIRTKLKAQREKASPOEH 1558
QY 1504 SAERTITIKAEYLSDDLKQSDSGVRSNESSPHSLHNEAADSQLEKANLIBLE 1563
Db 1559 SAERTITIKAEYLSDDLKQSDSGVRSNESSPHSLHNEAADSQLEKANLIBLE 1618
QY 1564 DEHSGKRGMPHSLSGLOPIIARMSICSEDKKSPSECSLIASSPESWPAQKAYNLNR 1623
Db 1619 DDSHSGKRGMPHSLSGLOPIIARMSICSEDKKSPSECSLIASSPESWPAQKAYNLNR 1678
QY 1624 TPSTVTLLNNTAPTNRANQNFDEIGIRETSQVILRPGPSPNPTAVQENILKSMHAKRSQ 1683
Db 1679 TPSTVTLLNNTAPTNRANQNFDEIGIRETSQVILRPGPSPNPTAVQENILKSMHAKRSQ 1738
QY 1684 RSVYTRLSKDASELH-AASSESTGGERESIL 1715
Db 1739 RSVYTRLSKDPPELHAAASSESTGGERESIL 1771

RESULT 13
ID ADC06850 standard; protein; 1142 AA.
XX AC ADC06850;
XX AC ADC06850;
DT 18-DEC-2003 (first entry)
DE Kidins 220Pc-related protein Pc473.
XX cytosolic; prostate cancer; breast; gene therapy; transgenic;
XX Kidins220Pc; kinase D-interacting substrate of 220KDa.
XX Unidentified.
XX WO2003064599-A2.

XX 07-AUG-2003.
XX 24-JAN-2003; 2003WO-US0001943.
XX 25-JAN-2002; 2002US-00054935.
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX WPI; 2003-679495/64.
XX New isolated polynucleotide related to cancer genes, useful for
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
PT or treating cancers, e.g. breast and prostate cancers.
XX Disclosure; Fig 20; 128pp; English.
XX The invention relates to a novel isolated polynucleotide comprising a
CC differentially-regulated mammalian cancer gene. The polynucleotides of
CC the invention demonstrate cytostatic activity and are differentially
CC expressed in prostate cancer. The polynucleotide, polypeptides and
CC methods of the invention may be useful for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating cancers,
CC particularly breast and prostate cancers. Furthermore, the invention may
CC be utilised during gene therapy procedures or in the production of
CC transgenic animals. The current sequence is that of the prostate cancer-
CC related protein of the invention. The C-terminus of the current sequence
CC is illegible within the specification and as a result, residues 1143-1715
CC are absent.
XX Sequence 1142 AA;
QY Query Match 64.9%; Score 5747; DB 7; Length 1142;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1097; Conservative 25; Mismatches 19; Indels 0; Gaps 0;
QY 1 MSVLISOSVINYVEENIPALKKLEKCKVDNERECQTPMLAAEQGNVEIVKELKN 60
Db 1 MSVLISOSVINYVEENIPALKKLEKCKVDNERECQTPMLAAEQGNVEIVKELKN 60
QY 61 GANCNLEDLNTWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGRDGV 120
Db 61 GANCNLEDLNTWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGRDGV 120
QY 121 ELLISHGANPSVTGLQSVYPIIWAAGRGHADIHLLONGAKVNCSDKYTTPLVWAAR 180
Db 121 ELLISHGANPSVTGLQSVYPIIWAAGRGHADIHLLONGAKVNCSDKYTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVDEQAGNSMTALIIVAVKGGYTQSVKEILKRNPNVLTDKGNTA 240
Db 181 KGHLECVKHLAMGADVDEQAGNSMTALIIVAVKGGYTQSVKEILKRNPNVLTDKGNTA 240
QY 241 LMIAKESGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMIAKESGHIIEVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
QY 301 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360
Db 301 QDNKTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360
QY 361 KVSADVKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
Db 361 KVSADVKKGTPLHVAIRGSRRLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480
Db 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480

```
Db 421 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFLLKKL 480
Qy 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVPAFPVDNLAIAISLSFLALIYIF 540
Db 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVPAFPVDNLAIAISLSFLALIYIF 540
Qy 541 FIVYFGRRGREGSNNWAWLSRLARHIGYELLLPKLMEFVNPPPELPEQTTKALPVRFPLF 600
Db 541 FIVYFGRRGREGSNNWAWLSRLARHIGYELLLPKLMEFVNPPPELPEQTTKALPVRFPLF 600
Qy 601 TDYRLSSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQKKKWKTCCLPS 660
Db 601 TDYRLSSVGGTSLAEMIAATLSDACEREFGLATRLFRVFRTEESQKKKWKTCCLPS 660
Qy 661 FVFLFVIGGIIAGITLLAIFRVDPKHLTVNAILISIASVVGAFVNLNCRTHWQVLDLSLL 720
Db 661 FVFLFVIGGIIAGITLLAIFRVDPKHLTVNAILISIASVVGAFVNLNCRTHWQVLDLSLL 720
Qy 721 NSQKRLHSAASKHLKSGFGMKVLCVELMARWAKTIDSTFQNTQRLVWIIDGLDAC 780
Db 721 NSQKRLHSAASKHLKSGFGMKVLCVELMARWAKTIDSTFQNTQRLVWIIDGLDAC 780
Qy 781 EQDKVLQMLDTVRVLFSGPFFIAIFASDPHIIKAINQNLNLSVLRDSNININGHDYMRNIVH 840
Db 781 EQDKVLQMLDTVRVLFSGPFFIAIFASDPHIIKAINQNLNLSVLRDSNININGHDYMRNIVH 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSOTGTQEDTDRVRSQNSLIGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSOTGTQEDTDRVRSQNSLIGEMTKLGSKTALN 900
Qy 901 RRDYRRRQORTITROMSFDLTKLVTEWDFSDISPDTRWRLNIVSVTGRLLRANQIT 960
Db 901 RRDYRRRQORTITROMSFDLTKLVTEWDFSDISPDTRWRLNIVSVTGRLLRANQIS 960
Qy 961 FNWDLASWINLQEQPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPPTTKDVEPLL 1020
Db 961 FNWDLASWINLQEQPYRTSWLILYLEETEGPLDQMTLKTMYERISKNIPPTTKDVEPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLECTVNLDPKLEIITADYRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLECTVNLDPKLEIITADYRAAREQINIGGLAY 1080
Qy 1081 PPLPLHGGPPRPSPGVSQPVASVCSSAFNGFPFGVVSPPHSSYSGLSGQHPFFYNRA 1140
Db 1081 PPLPLHGGPPRPSPGVSQPVASVCSSAFNGFPFGVVSPPHSSYSGLSGQHPFFYNRA 1140
Qy 1141 A 1141
Db 1141 S 1141

RESULT 14
ADCO6846
ID ADCO6846 standard; protein; 1184 AA.
AC
XX
AC ADCO6846;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human Kidney220Pc protein XM_045362.
XX
KW cytostatic; prostate cancer; breast; gene therapy; transgenic; human;
KW Kidney220Pc; kinase D-interacting substrate of 220KDa; chromosome 2p25.1.
XX
OS Homo sapiens.
XX
PN W02003064599-A2.
XX
PD 07-AUG-2003.
XX
PF 24-JAN-2003; 2003MO-US001943.
XX
PR 25-JAN-2002; 2002US-00054935.
```

```
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KP, Fan W, Shu Y;
XX
XX WPI; 2003-679495/64.
XX
PT New isolated polynucleotide related to cancer genes, useful for
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
PT or treating cancers, e.g. breast and prostate cancers.
XX
PS Disclosure; Fig 20; 128pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC differentially-regulated mammalian cancer gene. The polynucleotides of
CC the invention demonstrate cytostatic activity and are differentially
CC expressed in prostate cancer. The polynucleotide, polypeptides and
CC methods of the invention may be useful for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating cancers,
CC particularly breast and prostate cancers. Furthermore, the invention may
CC be utilised during gene therapy procedures or in the production of
CC transgenic animals. The current sequence is that of the prostate cancer-
CC related protein of the invention. The current sequence is not fully
CC legible within the specification; the complete sequence was obtained from
CC GenBank.
XX
SQ Sequence 1184 AA;
```

```
Query Match 64.7%; Score 5730.5; DB 7; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINVEENIPALKALLEKCKDVERNECGGTPLMAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINVEENIPALKALLEKCKDVERNECGGTPLMAAEQGNVEIVKELKN 60
Qy 61 GANCNLEDLNDNTALISASKEGHIHIVBELLKSGASLEHRDMGCTALMWACVKGRTDVV 120
Db 61 GANCNLEDLNDNTALISASKEGHVHIVBELLKCGVNLHRDMGCTALMWACVKGRTDVV 120
Qy 121 ELLLSHGANSPTGLQYVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLVAAAR 180
Db 121 ELLLSHGANSPTGL-QYVYPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPLVAAAR 179
Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 239
Qy 241 LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKVADIDIRG 300
Db 240 LMTASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKVADIDIRG 299
Qy 301 QDNKTALYWAVEKGNATVNRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVEKGNATVNRDILQCNPDTEICTKDGTEPLIKATKRNIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQKSL 420
Db 360 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNYIDCSHQKSL 419
Qy 421 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFLLKKL 480
Db 420 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFLLKKL 479
Qy 481 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVPAFPVDNLAIAISLSFLALIYIF 540
Db 480 EDEMKTFAGQOTPELPQFQSWLIIVFLTLILCGGLGLVPAFPVDNLAIAISLSFLALIYIF 539
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:18:20 ; Search time 40.5 Seconds
(without alignments)
3161.065 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSTGFGBERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pdp.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pdp.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pdp.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pdp.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pdp.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	575	6.5	2753	4	US-09-949-016-7659 Sequence 7659, Ap
2	575	6.5	2753	4	US-09-949-016-7660 Sequence 7660, Ap
3	558	6.3	1883	4	US-09-949-016-9010 Sequence 9010, Ap
4	558	6.3	1883	4	US-09-949-016-9011 Sequence 9011, Ap
5	558	6.3	1883	4	US-09-949-016-9012 Sequence 9012, Ap
6	558	6.3	1883	4	US-09-949-016-9013 Sequence 9013, Ap
7	558	6.3	1883	4	US-09-949-016-9014 Sequence 9014, Ap
8	558	6.3	1883	4	US-09-949-016-9015 Sequence 9015, Ap
9	558	6.3	1883	4	US-09-949-016-9016 Sequence 9016, Ap
10	558	6.3	1883	4	US-09-949-016-9017 Sequence 9017, Ap
11	556	6.3	1856	4	US-09-949-016-6964 Sequence 6964, Ap
12	556	6.3	1880	4	US-09-949-016-5876 Sequence 5876, Ap
13	556	6.3	1881	4	US-09-949-016-5965 Sequence 5965, Ap
14	550	6.2	3924	4	US-09-538-092-1246 Sequence 1246, Ap
15	545	6.2	1719	4	US-09-949-016-6966 Sequence 6966, Ap
16	536	6.1	1839	2	US-09-172-977-4 Sequence 4, Appli
17	536	6.1	1839	4	US-09-404-108-4 Sequence 4, Appli
18	525	5.9	4377	4	US-09-949-016-6978 Sequence 6978, Ap
19	503.5	5.7	843	2	US-09-172-977-3 Sequence 3, Appli
20	503.5	5.7	843	4	US-09-404-108-3 Sequence 3, Appli
21	466.5	5.3	1745	2	US-09-031-485-33 Sequence 33, Appl
22	466.5	5.3	1745	2	US-08-847-429A-33 Sequence 33, Appl
23	466.5	5.3	1745	3	US-09-065-474-33 Sequence 33, Appl
24	466.5	5.3	1745	3	US-09-557-034-33 Sequence 33, Appl
25	484.5	5.1	3913	4	US-09-949-016-10933 Sequence 2, Appli
26	449.5	5.1	1088	3	US-09-082-059-2 Sequence 8307, Ap
27	439.5	5.0	1881	4	US-09-949-016-8307 Sequence 8307, Ap

28	434	4.9	994	4	US-10-164-595-38 Sequence 38, Appl
29	414.5	4.7	1423	3	US-08-810-712-10 Sequence 10, Appl
30	414.5	4.7	1431	4	US-09-538-092-1198 Sequence 1198, Ap
31	409.5	4.6	683	4	US-10-164-595-71 Sequence 71, Appl
32	395.5	4.5	1327	3	US-09-196-387-2 Sequence 2, Appli
33	395.5	4.5	1327	4	US-09-841-835-2 Sequence 2, Appli
34	395.5	4.5	1327	4	US-09-972-115A-8 Sequence 8, Appli
35	394	4.5	1619	4	US-09-392-812A-4 Sequence 4, Appli
36	392	4.4	1704	4	US-09-392-812A-2 Sequence 2, Appli
37	387.5	4.4	949	3	US-09-196-387-10 Sequence 10, Appl
38	387.5	4.4	949	4	US-09-841-835-10 Sequence 10, Appl
39	384.5	4.3	546	4	US-09-949-016-7169 Sequence 7169, Ap
40	380.5	4.3	1100	4	US-09-427-154-2 Sequence 2, Appli
41	380.5	4.3	1100	4	US-09-696-668-3 Sequence 3, Appli
42	380.5	4.3	1166	4	US-09-972-115A-6 Sequence 6, Appli
43	380.5	4.3	1227	4	US-09-849-602-26 Sequence 26, Appl
44	380.5	4.3	1240	4	US-09-696-668-4 Sequence 4, Appli
45	376	4.2	348	2	US-09-031-485-28 Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-7659
; Sequence 7659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7659
; LENGTH: 2753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7659

Query Match 6.5%; Score 575; DB 4; Length 2753;
Best Local Similarity 21.7%; Pred. No. 2.5e-38;
Matches 411; Conservative 292; Mismatches 769; Indels 426; Gaps 76;
Qy 21 LKALLEKCKDVERNECGQTPMLAAEQGNVEIVKELKNGANCNLEDDNWTALISASK 80
Db 187 VELLLERGAFLARTKNGLSPLHMAAQGDHVECVKHLQHPVDDVTLDTLTAHVAAH 246
Qy 81 EGHITHIVEELKSGASLEHSDMGCTALMCKYKGRDVTVELLSHGNCAN-PSVTGLQVSV 139
Db 247 CGHYRVTKLLDKNRANPNARALNGFTPLHIAACKNRKIKVWELLVKYGASIOAIT--ESGL 304
Qy 140 YPIIWAAGRGHADIVHLLLLQNGKVCNSDKYGTTPLVAAARKGHLECVKHLAMGADVQ 199
Db 305 TPIHVAAFMGLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVRCCLLRNGALVDA 364
Qy 200 EGANSMTALIVAKGVGTQSVKEILKNPNVNLTKDGNLTALMTASKEGHIETIVQDILDA 259
Db 365 RAREEQTPPLHIAASRLGKTEIVQLLQHMAHPDAATNGYTPPLHISAREQVDVASVLEA 424
Qy 260 GTVYNIPDRSGDVTVLICAVRGHVEIVRALLQKADIDIRGQDNKTALYNAVEKGNATMV 319
Db 425 GAAHSLATKKGFTPLHVAAKYGLDVAKLQRRRAADSAGKGLTPLHVAHAYDNQKVA 484
Qy 320 RDILQCNPDTEICTKDGFTPLIKATKMRNIEVVVELLDKGAHSVADVKCKGDTPLHVAIRG 379

Best Local Similarity 20.6%; Pred. No. 3.5e-37;
Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;

QY	21	LKALLECKDQVDRNECGQTPLMLAAEQNVIEVKELLKNGANCNLEDLDNWTALISASK	80
Db	97	VELVNYGANVNAQSQKGFPLPYAAQENHLEVVFVLLGANGQNVATEDGFTPLAVALQ	156
QY	81	EGH-----	83
Db	157	QGHENVVAHLNYGKGVKVLPLALHIAARNDDTETAATVLLQNDPNPDVLSKTGFTPLHIA	216
QY	84	-----THIVEELKSGASLEHRDGMGTALMWACYKGRTDVVELLSHGANGPSVTGLQYS	138
Db	217	AHYENLNVAQLLNRGASVNFPTQNGITPLHIASRRGNVIMVRLLLDRGAQIE-TTKTDE	275
QY	139	VYPIIWAARGHADIHILLONGAKVNCSDKYGTTPLVMAARKGHLECVKHELLAMGADV	198
Db	276	LTPHCAARNGHVRISEILLDHGAPIOAKTNGLSPIHMAAQGDHLDVCRVLLLYQDAEID	335
QY	199	-----OEGA-----NSMTALIVAKGGVTSQVKILK	225
Db	336	DIPLDHLTPLVAAHCGHVRVAKVLLDKGAKPNSRALNGFTPLHIACKKHVRVWELLK	395
QY	226	RNPVNILTDKGNLTALMIASKEGHIETVOLLDAAGTYVNIPIDRSGDVTVLIVAGRGHVEI	285
Db	396	TGASIDAVTESGLTPLHVASFMGHLPVKNLLQRGASPNVSVKVEPLHMAARAGHTEV	455
QY	286	VRALLQKYADIDIRQDNKTALYVAVKGNATVMDILOCNPDTEICTKOGETPLIKATK	345
Db	456	AKVLLQNAKVNAKAKDQDTPLHCAARIGHTNMVKLLENANPNLATTAGHTPLHIAAR	515
QY	346	MRIEVEVELLDKGAQSAVDKGDTPLVHVAIRGSRRLABELLNRNPKDGLLRPNKAG	405
Db	516	EGHVETVIALEKEASQCMTKGFTPLHVAAKYGVKRVAEILLLE--RDA-----HPNAG	569
QY	406	E--TPYNIIDCSHO--KSLI-----TQIFGAR-----	427
Db	570	KNGITPLHVAVHNNLNDIVKLLPRGGSPHSPAWNGYTPLHIAAKQNEVARSILQYGG	629
QY	428	-----HLSPTETDGMGLVDLYSSALADILSEPTMP-----PIC--	462
Db	630	SANAESVQGVTPPLHAAQEGHAEVALLSKOANGNLGNSGLTPLHLVAGEHVPVADV	689
QY	463	-----VGLYAQWGSCK--SFLKKLED-EMKTFAGQOETPLFQFSWL	501
Db	690	LIKHGVMDATRMGYTPLHVASHYGNIKLVKFLLOHQADVNAKTKLV--SPLHQAAQQ	747
QY	502	--IVFLTLLCGGJGLVPAPVD-----TNLAIAISLSFALYIPIFVIVYFGRREGE	553
Db	748	GHTDITVLLKNG-----ASPNEVSSDGGTTPLAIAKRLGYISVTDVLKVW-----	792
QY	554	SWNWAWALSTLASHIGYLELLFKLMFVNPELPEQTTKALPVRFLETDYNRLS-SVGGE	612
Db	793	TDSTSFVLVSDKHR--MSFPVTDEILDVSEDEGBELIS-----FKAERDRSDVDEE	843
QY	613	TSLAEMIATLSDACEREGFLATLFR-----VERTBE-SQKKKKWTKTCLCPFV	662
Db	844	KELLDVFPKLQDVE-----SPALPRPCAMPETVIVIRSEEQASKEYDEDSLIPS--	895
QY	663	IFLFTVGCIIAGITLLAIFRVDPKHLTVNAILISIASVVGLAFLVNLNCTRWQVLDLSLNS	722
Db	896	-----SPATETSDNI--SPVASPVHTGFLVSF-----MVDARGGS	928
QY	723	QRKRLHSAASKHLKLSGFKMVKCEVELMARMAKTIDSTFQNTQTRLVVIIDGLDACEQ	782
Db	929	MRGSRHN-----GLRVIP-----PRTCAAPTRITCRLV-----	957
QY	783	DKVLQMLDTRVLFPSKGPPIAFASDPHIIKAINONLNSVLNRDSNINGHDYWRN-IVHL	841
Db	958	--KPQKLSPPPL-----AEEGLASRIIALGPTGAQPLSPVIVEI	996
QY	842	PVFL-NSRGLSNARKFLVTSANGDITCSDDTTGTQEDTDRRVQNSLGMTKLGSKTALN	900

RESULT 4

RESULT 4
US-09-949-016-9011

US-09-349-016-9011
: Sequence 9011, Application US/09949016

Sequence 3011, App. 12339
Patent No. 6812339

; Patient NO. 6812332
; GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

APPLICANT: VENTER, O. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN CLASSES OF COMPOUNDS
 ;
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; TITLE OF INVENTION: WITH HUMAN DISEASE;
: FILE REFERENCE: CL001307

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/944

;; CURRENT APPLICATION NUMBER: US/09/545,018
: CURRENT FILING DATE: 2000-04-14

; CURRENT FILING DATE: 2000-04
; PRIOR APPLICATION NUMBER: 60/

Ddb		997	PHFASHGRG----	DRELVVRSENGSV-----	WKEHSRYSGSYLDQILN-GMDEELG	1044
Qy		901	RRDTYRROQMORTITROMSFO	LTKLVLT--EDWFSDISPTMRRLLNIVSVTGRLLRAN	957	
Ddb		1045	SLEELEKRVCKRIIT--TD	FPFLYFVINSRLCOD-YDTIGPEG-----	GSLKSKLPLVLV	1094
Qy		958	QITFNWDRLASWINLTEOWPRT	SWLLIYLBEETEGCLPDOMTLK-----	TMYER	1005
Ddb		1095	QATPPENAVTRVKVALQ-		AQVPDELVTYKLIGNOATTSPIVTVPR	1140
Qy		1006	ISK-----	NIPTTKDVEFLLEIDGDIRNFVF-----	LSSRTFVLVA	1042
Ddb		1141	PKKHPRPICLRIPIPSPWTDP	RDSGEGETTSLRLCSVIGGTDOAQWEDITGTTKLVA	1200	
Qy		1043	RDVKTFTCTVNLPDKREIIAD	VRAAREQINIGGLAYPPILPHGEGPPRPSPGISQPASV	1102	
Ddb		1201	NECANF---TTNV-SARF	WLSDCPRTAEVAFATLYKEL-----	TAVPYMAKF	1245
Qy		1103	CSSASFNGCFPGGVSPQHPS	SYTSGLSGPQHPFYNNRAAVPATGSLLLSSMTVDVWCEK	1162	
Ddb		1246	VIFAKNDPREG-----		RLRCYCMTDDKVDT	1272
Qy		1163	LRQ-----	IEG-----LDONMPOYCTTIKAN-----	INGRVL	1191
Ddb		1273	LEQHENVFEVARSRDIEVL	EGMSFAELSGLNLVP----VKKAAQRGFHQFSRENRLA	1327	
Qy		1192	SOCNIDEUKZE-----	MAMFGD-WHLFRSMVLEMRSVESQVVEDPR-----	1233	
Ddb		1328	MPVKVRDSRSREGGSLSF	LRKAMYEDTQHILCHLINITMPPCAKGSGAEDRRTTPPAL	1387	
Qy		1234	---FLNENSSAPVPHGESARR---	SSTHELPLTELSSQTPTYTLNFSPHEELNTLGLD---	1283	
Ddb		1388	RYSIUSESTPGSLSTEQAEM	KMAVISHEHGLSWAELARE---LQFSVEDINRIVENP	1443	
Qy		1284	----EGAPRHNSLWSQSTR	TPSLSSLNSSODSSTEISKLTKVOA-----EYRD	1329	
Ddb		1444	NELLEQSALLNL-WVIRE	GOANNENLYTALQSIDRGIEVNMLEGSGRQSNLKPDREH	1502	
Qy		1330	AYREVIAQMSOLEGGTSG	STISGRSPHSTYYIGOSSGSHSTLEQBRGEGELKOED	1389	
Ddb		1503	TDRTDSLSPSQNG--YSSL	QODELLSP-----ASLGALUSSPLRADQ-----	1542	
Qy		1390	GRKSFLMKRGDVIDYSSG	SVSTNEASPDLPI---TEED---EKSDQS--GSKLLPGKKS	1440	
Ddb		1543	-----	YNNEVAVILDAIPLAATEHDHTMLEMSDMQVMSAGLTPLSVT	1582	
Qy		1441	SERPSLFQTDCLKGGGLRY	KQLPDEDESGETGRVQITPHCSKMINTKELKAKORCASP	1500	
Ddb		1583	AEDSSL-----	ECSKAEDSDATGH-----EWKLEGAUSEEPGRP	1616	
Qy		1501	QEHSAREPIRTFKAKEYL	SDA---LLDKKDSGDSGVRSNESSPNHSLHNAAADDSOLEKA	1557	
Ddb		1617	ELGSLB-----	LVEDDTVDSDATNGLIDLLE-QEEGORSEEKLPGSKRQDDATGAGQ-DSE	1670	
Qy		1558	NLIELEDEHGSGKRGMPH	SLSGLODPIIARMICSEDK-----KSPSECSLIASPEESW	1612	
Ddb		1671	NEWSIVSGHORGOARTHS-----	PTVSQVTERSODRLQDMDADGSGIVSYLQDAQQSGW	1724	

QY 1330 AYREYIAQMSOLRGSTTSIGRSPSHSTYIIGQSSSGSIHSTLEQERKGEKELQED 1389
Db 1503 TDRDYSLSPSQMG--YSSIQDELLSP-----ASLGCALSSPLRADO----- 1542
QY 1390 GRKSFLMKRGDVIDYSSGVSTNEASPLDPI-----TEED---EKSQDS--GSKLLPKKKS 1440
Db 1543 -----YNEVAVLDAIPLAAYTEHDTMLEMSDMQVWSAGLTSLVT 1582
QY 1441 SERPSLFTDLKKGGLRYOKLPSEDESGTRVQIITHCSKMIRTKRLKAKORECASP 1500
Db 1583 AEDSSL-----ECSKAEDSDATGH-----EWKLEGALSSEPRGP 1616
QY 1501 QEHSAEPIRTFIKAKEVLSDA---LLDKDSSDGSVRSNESSPNHSLHNEAADDSOLEKA 1557
Db 1617 ELGSLE-----LVBDTVDVSDATNGLIDLE- QEGORSEBKLPGRQDDATGAGO-DSE 1670
QY 1558 NLIELEDEGHSGRGMPSHLSGLQDPIIARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1671 NEVSLVSGHORGQARITHS-----PTVSQVTERSQDLQDWDADGSIYSYLQDAAGSW 1724

RESULT 7

US-09-949-016-9014
; Sequence 9014, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9014
; LENGTH: 1883
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9014

Query Match 6.3%; Score 558; DB 4; Length 1883;
Best Local Similarity 20.6%; Pred. No. 3.5e-37;
Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;
QY 21 LKALLECKDVERNECGQTPLMLAABQGNVEIVKELLKNGANCNLEDLNNWTALISASK 80
Db 97 VRELIVYGANVNAQSGKFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQ 156
QY 81 EGH----- 83
Db 157 QGHENVVAHLINYGTKGKVRPLPALHIAARNDDTTRTAIVLQNDPNPVLKGTGTPHIA 216
QY 84 -----IHIVEBLLKSGALBHRDMGWTALMWACYKGRDVTVDVLLSHGANPSVTGLQYS 138
Db 217 AHYENLVNAQLLNRGASVNFPTQGITPLHIASSRGVIMVRLLLDRGAQIE-TTKDE 275
QY 139 VYPIIWAAGRHADIHLLONGAKVNCSDYKGTTPLVWAARKHLECVKHLLAMGADVD 198
Db 276 LTPLHCAARNGHVRISILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVCRLLQLQYDAID 335
QY 199 -----QEGA-----NSMTALIVAVKGGVTSVKELK 225
Db 336 DITLDHUTPLHVAACHHVRVAKVLLDKGAKPNRSALNGFTPLHIACKKHVRWMLLLK 395
QY 226 RNPVNLTDKGNLTALMIASKEGHIEIVQDILLDAGTVYINIPDRSGDVTVLIGAVRGHVEI 285

Db 396 TGASIDAVTESGLTPLHVASFMGHLPIVKNLQQRGASPNVSVKVTPLHMAARAGHTEV 455
QY 286 VVALLOKYADIIRGDNKNTALYWAKEGNATWVRDILQCNPDTEICTKDGTEPLIKATK 345
Db 456 AKYLLQNKAKVNAKDDQDTPLHCAARIQHTNMVKLLLENNANPNLATTAGHTPLHIAAR 515
QY 346 MRNIEVVELLLDKGAKVSAVDKGGTPTPLHVAIRGRSRRRLAEILLRPNKGRLLRYRPNKAG 405
Db 516 EGHVEIVLALLKEASQACWKTKGFTPLHVAAYKGVVRAELLLE--RDA---HFNAAAG 569
QY 406 E---TPYNTDCSHQ-----KSIL-----TQIFGAR----- 427
Db 570 KNGLTPLHVAHNNLDIVKLLPRGSGSPHSPAWNCGYTPHIAAKQNVARSLLQYGG 629
QY 428 -----HLSPTETDGMGLGYDIYSSALADILSEPTMQP-----PIC-- 462
Db 630 SANAESVOGVTPLHLAAQEGHAEWALLLSKQANGLNKSGGLTPLHLVAQEGHVPADV 689
QY 463 -----VGLYAQMGSGK--SPLKLEJED-EMKTFAGQOTTEPLFQFWSL 501
Db 690 LIKHGVMVDATTMGVYTPHVAASHYGNIKLVKFLQHQADVNAKTKLGY--SPLHQAAQ 747
QY 502 --IVFLTLLCGGLGLVFAFPVD-----TNLAISLSFLALIYIFFIVIFVGGREGE 553
Db 748 GHTDIVTLLKNG-----ASPNEVSDGTPLAIAKRLGYISVTDVLKV- 792
QY 554 SNWAWALSTRLARHIGYLELLPKLMFVNPPPEQTTRKALPVRFLFTDYNRLS-SVGGE 612
Db 793 TDTSFVLVSDKXR-MSFPETVEILDVSEDEGEELIS-----FKARRDSRDVDEE 843
QY 613 TSAEMIAIATLSACEREFGLATRLFR-----VFRTEE-SQGGKKWKTCCLPSFV 662
Db 844 KELDFVPKLDQVVE-----SPAIPRIPCAMPETVIRSEEQEQAKEYDEDSLIPS-- 895
QY 663 IFLFIVGCIAGITLAI FRVDPKHLTVNAILISIASVGLAFVLCNRTWQVLDSELS 722
Db 896 -----SPATETSDNI-SPVASPVHTGTGLVSF-----MVDARGGS 928
QY 723 QRKRLHSAASKLHLKSEGMKVLKCEVELMARMAKTIDSTQNTQRLVVIIDGLDACEQ 782
Db 929 MRGSRHN-----GLRVVIP-----PRYCAAPRITCLV----- 957
QY 783 DKVLQMLDTRVRLFSKGPFIAPASDPHIIKAINQNLNSVLNRDSNINGHDYMRN-IVHL 841
Db 958 --KPQKLSPPPL-----AEEGLASRIIALGPTGAQFLSPVIVEI 996
QY 842 PVFL-NSRGLSNARKFLVTSANGDITCSDTGTQEDTDRRSQNSLIGEMTKLGSKTALN 900
Db 997 PHFASHGRG---DRELVLRSNGSV-----WKEHRSRYGESYLDQILN-GMDEELG 1044
QY 901 RRDYTRRRQWORTITRQMSFDLTCLKAVT---EDWFSDDISQPTWRRLLNIVSVTGRLLRAN 957
Db 1045 SLEELKRVCRIT--TFPLFYVIMSLCQD-YDTIGPEG-----GSLAKSLVPLV 1094
QY 958 QITFNDRILASWINLTEQMPYRTSWLILYLEETEGPLDQMTLK-----TWYER 1005
Db 1095 QAITFPENAVTKRVKLAQ-----AQVPDELVTCLKGNOATFSPVITVEPR 1140
QY 1006 ISK-----NIPITKQVPELLEIDGDIRNEFV-----LSRTPVLVA 1042
Db 1141 RRKPHRPIGLRIPPLPSWTDNPRDSGEGDTLSRLLCVSGVIGGTDQAQWEDITGTTKLVA 1200
QY 1043 RDVKTFLPCVTNLDPKLREIADVRAAREQINIGLAYPPLPLHEGPPRPPSGYSPASV 1102
Db 1201 NECANF--TTNV--SARFWSLDCPRTAAVNFATLLYKEL-----TAVPYMAKF 1245
QY 1103 CSSASFNPGPPGVWSQPSSHYSYGLSGPHFPYNAAVPATGSSLLSSMTVDVVECK 1162
Db 1246 VIFAKNDPREG-----RRCYCMTDDKVDKT 1272
QY 1163 LRQ-----TEG-----LDQNMMPQYCTTIKKAN-----INGRVL 1191
Db 1273 LEQHENFVEVARSRDIEVLGMSLFAELSGNLVFP-----VKAAQORSFHQSFQFRENRLA 1327


```
QY 428 -----HLSPETDGDMLGYDLYSSALADILSEPTWQP-----PIC-- 462
Db 627 SANAESVOGVTPLHLAAQEGHAEMVALLSKQANGNIGNKSGLTPLHLVAQEGHPVADV 686
QY 463 -----VGLYAQWGSCK--SFLKKKLED-EMKTAGQOQTEPLFPFQSWL 501
Db 687 LIKHGVAVDATTMGVYPLHVASHYGNIKLVKFLHQADVNAKTKLGY--SPLHQAQQ 744
QY 502 --IVFTLLCGGLGLVAFPD-----TNLAIAISLFLAIYIFVIFYPGRRGE 553
Db 745 GHTDIVTLLKNG-----ASPNEVSSDGTTPLAIAKRLGYSIVTDVLKV- 789
QY 554 SNKAWALSRLARHICYELLFKLMFVNPPELPEQTTKALPVRFUFDYNRLS-SVGGF 612
Db 790 TDTSFVLVSDKRR-MSFPETVDEILDVSEDEGELIS-----FKARRDRSDVEE 840
QY 613 TSAEMIATLSDACERFGLFATLRF-----VFRTEE-SQKKKKWKTCCLPSPFV 662
Db 841 KELLDFVPKLDQVVE-----SPAIPRIPCAMPETVIRSEEQEASKEYDEDSLIPS-- 892
QY 663 IFLFVGCIIAGITLLAIFRVDPRKHLTVNAILISIASVGLAFVNLNRTWMQVLDLSLNS 722
Db 893 -----SPATETSDNI-SPVASPVHTGFLVSP-----MVDARGGS 925
QY 723 QKRLHSAASKLHLKSEGFWMKVLKCEVELMARWAKTIDSTQNTQTLVLIIDGLDACEQ 782
Db 926 MRGSRHN-----GLRVIP-----PRTCAAPTITRLV----- 954
QY 783 DKVLQMLDVRVLFSGPFIAPASDPHIIKAINONLNSVLDRSNNINGHDYMRN-IVHL 841
Db 955 --XPQKLSPPPL-----ABEGLASRIIAGTGAQFISPVIVEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTTGQEDTDRRVSONSGEMTKLGSKTALN 900
Db 994 PHFASHGRG--DRELVLRSNGSV-----WKEHRSRYGESYLDQILN-GMBEELG 1041
QY 901 RRTYRRQORITROMSFDLTKLVT-----EDWFSDISPQWRRLLNIVSVTGLLRAN 957
Db 1042 SLEBEKKKRCRIIT--TDPLFYVIMSLRQCD-YDIIGPEG-----GSLKSLVPLV 1091
QY 958 QITFNMDRLASWINLTQWVPYRTSWLLIYLEETEGPLDQMTLK-----TWYER 1005
Db 1092 QATFEPENAVTKRVKALQ-----AQVPDELVTKLGNQATFSPVITVEPR 1137
QY 1006 ISK-----NIPTKKOVEPLEIDGDIRNEFV-----LSSRTPVLVA 1042
Db 1138 RRFPHRPIGLRIPLPSSWTNPRDSGEGDTLSRLLCVIGGTDQAQWEDITGTIKLVA 1197
QY 1043 RDVKTFLPCTVNLDPKLEIIADVRAREQINIGGLAYPPLPLHEGPPRPPSGYQPASV 1102
Db 1198 NECANF---TTNV--SARFWLSDCPRFAEAVNFATLLYKEL-----TAVPYWAKF 1242
QY 1103 CSSASFNPGPPGGVSPQPHSSYVSLSGPQHPFNRAAVPATGSSLLSSMTVDVVCCK 1162
Db 1243 VIFAKNDPREG-----RLRCYCMTDDKDKT 1269
QY 1163 LRQ-----IFG-----LDQNMMPQYCTTIKKAN-----INGRVL 1191
Db 1270 LEQHENFEVARSDIEVLGMSLFAELSGNLVP-----VKAAQQRSFHFQSFRENRLA 1324
QY 1192 SQCNIDELKYE-----MANNFGD-WHLFRSMVLEMRSVESQVPEPDR----- 1233
Db 1325 MPVKVRDSSRPGGSLFKAKMYEDTQHILCHLNITMPPCAKGSAGARRRTPPTPLAL 1384
QY 1234 ---FLNNSAPVPHGSAAR-----SSHTLPLTELSQTPYTLNFSFEELATGLD-- 1283
Db 1385 RYSLSESTPGSLGTEQAEWKAVISEHLGLSWAELARE---LQFSVEDINRIVENP 1440
QY 1284 ----EGAPRHSNLSWOSQTRTPSLSSLSNSQDSSEISIKLTDKVA-----EYRD 1329
Db 1441 NSLLEQSVALLNL-WVIREQANMENLYTALQSIDRGEIVNMLEGSGRSLNKDPRRH 1499
```

1330 AVREYIAQMSOLEGGTSGSTISGRSSPHSTYYIYGSSSGSIHSTLEQERKEGELKQED 1389
1500 TRDYSLSFQNG--YSSLQDELLSP-----ASLGCALSSPLRADO----- 1539
1390 GRKSFMLKRGDVIDYSSSGSVSTNEASPLDP-----TEED---EKSDQS--GSKLLFGKKS 1440
1540 -----YNNVAVLDAIPLAATEHDTMLEMSDMQVMSAGLTPTSLVT 1579
1441 SERPSLFQTDLKLKGGGLRYQKLPSEDESGETGRVOITHCCKMIRTKRLKAKQRECAP 1500
1580 AEDSSL-----ECSKAEDSDATGH-----EMKLEGALSEPRGP 1613
1501 QEHSASPIRTFIKAKBYLSDA---LLDKXSSDSGVRSNESSPNHSLHNEAADDSOLEKA 1557
1614 ELGSLB---LVEDDTVDSATNGLIDILLE-OEGQORSEKLPGRQDDATGAGQ-DSE 1667
1558 NLIEDEGHSGKRGMPHSLSLGLQDPIIARMSICSEDK-----KSPSECSLIASSPEESW 1612
1668 NEVSLVSGHQGQARITHS-----PTVQVTERSQRLQDMDADGSIVSYLQDAAGGSW 1721

RESULT 13
US-09-949-016-6965
; Sequence 6965, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6965
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6965

Query Match 6.3%; Score 556; DB 4; Length 1881;
Best Local Similarity 20.6%; Pred. No. 5.2e-37;
Matches 395; Conservative 270; Mismatches 635; Indels 620; Gaps 75;

QY 21 LKALLEKCKDVBERNECGOTPLMLAAEQGNVEIVKELLKNGANCNLEDLNDWTALISASK 80
Db 94 VRELNYGANVNAQSQGFTPLYMAQENHLEVYKELLENGANQNVATEDGFTPLVALQ 153
QY 81 EGH----- 83
Db 154 QGHENVVAHLINYTKGVRLPALHIAARNDDTRTAAVLQNDPNPDVLSKTGFTPLHIA 213
QY 84 -----IHIIVEELKSGASLEHRDMGWTALMWACYKGRDVTVVELLSHGANPSVTGLQYS 138
Db 214 AHYENLNVALLNRCGASVFTPGNGITPLHIASRRGNVIMVRLLDRAQIB-TTKDE 272
QY 139 VPTIWAAGRHADI VHLILONGAKVNCSDKVTGTTPLVWAARKHLECUKHLAMGADV 198
Db 273 LTPHCAARNGHVRISEILLDHGAPIQATKNGLSPIHMAAQGDHLDCCVRLLOYDAEID 332
QY 199 -----OEQA-----NSMTLIVAVKGYTOSVKEILK 225
Db 333 DITLDHLTPLHVAHCHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKHVVRVMEILLK 392
QY 226 RNPVNLTKDGNLTALMIASKEGHIIEVDLLDAGTYVNIIPDRSGDTVLIGVRGGHVEI 285
Db 393 TGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSVKVVETPLHMAARAGHTEV 452

133 PLYMAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQHQVAJLLENDTKGKVR 192
104 -----GWTALMAWCKGRTDVELL 123
193 LPALHIAARKDDTKSAALLQNDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLL 252
124 LSHGANPSVTGLQSVVPIIWAAGRHADIVHLLQNGAKVNCSDKVGTTPLVWAARKG- 182
125 LNRGAADVFTA-RNGITPLHVASKRGNTNWKLLDRGGQIDAKTRDGLTPLHCAARSGH 311
183 -----HLECVKHLAMGADVDOEGANSMTALIV 210
312 DQVVELLERGA PLLARTKNGLSPLHMAAQGDHVECVHLLQHKAPVDDVTLDTALH 371
211 AVGGYTQSVEIKRNPVNLTKDGNALMTASKE----- 247
372 AACHGHRVTKLLDRKANPNARALANGFTPLHACKNRKIKVMELLVKYGASIQAITESG 431
248 -----GHTEIIVQDLDADGTYVNI PDMSGDTVLIGAVRGHVEIVRALLQKVADID 297
432 LTPHVAAFMGHLNIVLLLLQNGASPDVWIRGETALHMAARAGQVEVRCLENGALVD 491
298 IRQDNKTALYAWVEKGNATWVRDILQCNPDTEICTKGETPLIKATKMRNIEVVELLD 357
492 ABAREETPLHIASRLGKTEIVQLLQHMHPDAATNGYTPLHISAREGOVDVASVLE 551
358 KGAKVSAVDKGTPLHVAIRGRSRLAELLR-----NPKD 394
552 AGAAHSLATKGTGFTPLHVAAYKGSVDVAKLLQRRAAADSAGKNGLTPLHVAAHYDNQKV 611
395 GRLL-----YRPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPT 432
612 ALLLEKGA PHATAKNGYTPLHIAAKNQWQASTLLNGAETNIVTKGVTPHLASQ 671
433 ETQDML-----GYDIYALADILSEPTMOPPICVGLYAQMGSGKSFLLKKLEDEM-- 484
672 EGHTDMVTLLLDKGNIHMTSGSLTS-----LHLAAQ-----EDKVV 710
485 -----KTFACQOETPLPQFWLIVFLTLILC--GGGLV-FAPVDNLAIALSLSPAL 536
711 ADILTRHGAQDAHTKGLGYPLIV-----ACHYGNVQVNFLLKQGANVNAKTKNGYTPL 765
537 -----IVIFFIIVFGRRGESMNAWALSTRLARHIGYLELLFKLMFNPNP----- 584
766 HQAAQQGHTHILVLLQHGAKPNATTANGNTALA--IAKGLVSVVDLTKVTEEVTTT 823
585 -----ELPEQTTKALPVRFELTDYNRLSVSGGETSLAEMIAFLSDACREBFGLA 634
824 TTTITEKHLNVPTMTVEVLVDVDEGD-DTMTGDGGEYLRPEDLKELGDDSLPSSQFLD 882
635 TRLFRVFRTEESQKKWKTCCCLPSFVIFLIVGCIAGITLAIAPRVPKHLTVNAIL 694
883 GMYLRYSLGGRSDS-----LRSPSDRSHTLSHASY 915
695 ISIASVVGAPVL-----NCKTW--QVLDLNLNQRKRLHSAASKLHL 737
916 LRDSAVDDSVVISHQVSTLAKAERNVRLSNGTENLNVALS-----SSPH-- 965
738 KSEGMKVLKCEVELMARMARKTIDFTQNTQNLVVIIDGLDACQDKVLQMLDTRVLF 797
966 --SGFLVIFMVDAR-----GGMRGCRHNGLR--IIIPRKTAPTTRVTCRLVKRHLAT 1016
798 KGPFIAPASPHIILKAINQNLNVLNDSNINGHYEN-IVHLPVFLNLSRGLSNARKF 856
1017 MPFVVE-----GEGLASRLIEVFGSAQFGLPVIVPHFAALRG--KREL 1061
857 LVTGATNGDI-----TCSDTTGTQBDTRRSQNSLGEWTKLGSALTNRDTRRRQMR 912
1062 VLRSENGDSWKEHFCDYT-----EDELNEILN-GMDEVLDSPDELEKRICR 1108
913 TITROMGFDLTKLVTEDMFSDISQTNWRLNLTIVSVTGRLLRAN-----QITFNWDLA 967
1109 IITRDF-----POYFAVVS--RIKQDNLNIGPEGGLVSTVVPQVAFPEGALT 1156

QY 968 SMINL-TEOWPYRTSWLILYLEETEGLPDQMTL-----KTMYERISKNIPTTK--DVEPL 1019
Db 1157 KRIRVGLQAOQPMHSELVKILGNKATFSPIVLEPRRRKFKHPITWTIPVPRASDV-ML 1215
QY 1020 LEIDGDIRNEVF-----LSSRTPVLVARDVTKTFLPCTVNLDPKLRILIID 1065
Db 1216 NGFGDAPTRLRLCSITGTTTQAQWEDITGTTPLTFVNECVSF--TTNV--SAREWLD 1270
QY 1066 VRAAREQINIGLAPPLPLHEGPPRPSPGYSQSPASVCSASFPNGPFGVVSPOPHSSV 1125
Db 1271 CRQIQESVTFASQVREI-----IC-----VPMK 1296
QY 1126 YSGLSQPQHPFNRAAVPATGSLSSMTVDVCEKLRQIEGLDQNMMPQYCTTIKKAN 1185
Db 1297 FVFAKSHDPIEAR-----LRCFCMTDDKDKTLEQ-----QENFAVARSDEV 1342
QY 1186 INGR-VLSQC--NIDELKEMAMNFGDWHLFMSVLEMRSVSOVVPDPFRLENSSAP 1242
Db 1343 LEGKPIYVDFCNLVP LTKS-----GOHIFSFAPK-----ENRL-----PLFKVRDITQ 1389
QY 1243 VPHGSARSSHTPELPTLSSTQPTYNFSEELNTLGLDEGAPHNSL----- 1292
Db 1390 EPCGR-----LSFMKEPSTRGLVHOAILNLTITPIYTKESB 1427
QY 1293 SWQSOTRRTPSSLNSQDSSIEISKLTDKVQAEYRD-AYRYVIAQMSOLEGSGTGSSTIS 1351
Db 1428 SQEQEEEDMTSEKNDTESTETSVLKSHLVNEVPVLASPDLLSEVSEMQLDIKWT-- 1485
QY 1352 GRSSPHSTYIIOSSSGGSIH-----STLEQBRGEGELKQEDGRKSFMLKGDVDIDYS 1406
Db 1486 -----AILTVDSDKAGSIKVELVKAABEFPGEFIVER--VKEDLEKVNEL--R 1534
QY 1407 SGVSTNEASPLDP-----ITEE-----DEKSDQSGKL-----LPG 1437
Db 1535 SGCTRDESSVQSSRSERGLVEEWVIVSDELEEARQKAPLEITPYCVVEVRIDKEIG 1594
QY 1438 KKSSERP SL--FOTD-----LKLKGG-----GLRYOK-----LPSDED 1468
Db 1595 KVEKSTGLVNYLTDDLNTCVPLPKELQTVQDKAGKCEALAVGRSSEKGDIPDET 1654
QY 1469 ESGTGRVQITPHCSKMRIRKAKORECASPOEHSABEIRFIKAKELYSDALLDKDS 1528
Db 1655 QSTQKHQPSGLIGKFPVR-RKLKEKQKEGLQASAE-----KAE-----KKS 1699
QY 1529 SDSGVSNSSPNHSLHNEAADDQLE-----KANLIEDEGHSGKGMPSHL 1577
Db 1700 SEESGEDPGLAPEPLTVKATSPLEETPIGSIKDKVAKLQKVEDE-QKGRSKLPIRV 1758
QY 1578 SGLQDPIIARMSICSBDKSPSECSLIASSPEESWPACQKAYNLRNTPSTVTLNNNTAPT 1637
Db 1759 KGKED-----VPKKTTHRPHPAAS-----PSLKSERHAPGSPKTERHSTLS 1802
QY 1638 NRANQFDEIEGIRETSQVILRPGSP-NPTAVQNONELKSMHAKRSQSRSYTSLKDAE 1696
Db 1803 SAKTERHPVPSPSKTEK-----HSPVSPSA-----KTERHSPASSSKT-----E 1843
QY 1697 LHAASSEST 1705
Db 1844 KHSFVSPST 1852

RESULT 15

US-09-949-016-6966
; Sequence 6966, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:28:23 ; Search time 141 Seconds
(without alignments)
3961.558 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSTGFGBERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:**
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8853	100.0	1715	14	US-10-021-571-2
2	8853	100.0	1715	14	US-10-117-229-6
3	8639.5	97.6	1762	14	US-10-117-229-7
4	8266	93.4	1715	14	US-10-021-571-4
5	8266	93.4	1715	14	US-10-117-229-2
6	8266	93.4	1715	16	US-10-473-574-26
7	8212	92.8	1771	14	US-10-117-229-4
8	8212	92.8	1771	14	US-10-117-229-11
9	5730.5	64.7	1184	14	US-10-117-229-3
10	5730.5	64.7	1184	14	US-10-117-229-9
11	2961.5	33.5	705	14	US-10-106-698-6378
12	2782.5	31.4	551	9	US-09-835-788A-17
13	2782.5	31.4	551	14	US-10-175-042-17

14	2376.5	26.8	1498	14	US-10-021-571-8	Sequence 8, Appli
15	2260	25.5	513	14	US-10-149-819-9	Sequence 9, Appli
16	1848.5	20.9	1398	14	US-10-021-571-6	Sequence 6, Appli
17	549	6.2	109	10	US-09-986-480-371	Sequence 371, App
18	538.5	6.1	1724	9	US-09-964-899-43	Sequence 43, Appli
19	527.5	6.0	657	15	US-10-104-047-2529	Sequence 2529, Ap
20	506.5	5.7	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
21	483.5	5.5	1762	14	US-10-205-194-117	Sequence 117, App
22	479.5	5.4	1188	16	US-10-408-765A-1311	Sequence 1311, Ap
23	456	5.2	747	15	US-10-094-749-1924	Sequence 1924, Ap
24	454.5	5.1	3913	15	US-10-334-143-45	Sequence 45, Appli
25	438	4.9	740	9	US-09-835-788A-12	Sequence 12, Appli
26	438	4.9	740	14	US-10-175-042-12	Sequence 31, Appli
27	438	4.9	1569	15	US-10-275-595A-31	Sequence 4122, Ap
28	434.5	4.9	919	15	US-10-108-260A-4122	Sequence 343, App
29	431.5	4.9	1053	15	US-10-291-172-343	Sequence 343, App
30	431.5	4.9	1053	15	US-10-221-278-343	Sequence 38, Appli
31	414.5	4.7	1431	15	US-10-352-684A-38	Sequence 5, Appli
32	414.5	4.7	1431	15	US-10-045-400C-5	Sequence 1235, Ap
33	414.5	4.7	1431	15	US-10-295-027-1235	Sequence 70, Appli
34	401.5	4.5	426	9	US-09-908-711-70	Sequence 6, Appli
35	400	4.5	367	15	US-10-250-613-6	Sequence 2, Appli
36	395.5	4.5	1327	9	US-09-841-835-2	Sequence 8, Appli
37	395.5	4.5	1327	10	US-09-972-115A-8	Sequence 4, Appli
38	395.5	4.5	1327	14	US-10-199-937-4	Sequence 3, Appli
39	394.5	4.5	765	14	US-10-128-174-3	Sequence 34, Appli
40	394.5	4.5	765	14	US-10-128-174-34	Sequence 35, Appli
41	394.5	4.5	765	14	US-10-128-174-35	Sequence 36, Appli
42	394.5	4.5	765	14	US-10-128-174-36	Sequence 37, Appli
43	394.5	4.5	765	14	US-10-128-174-37	Sequence 38, Appli
44	394.5	4.5	765	14	US-10-128-174-38	Sequence 39, Appli
45	394.5	4.5	765	14	US-10-128-174-39	

ALIGNMENTS

RESULT 1
US-10-021-571-2
; Sequence 2, Application US/10021571
; Publication No. US20030168056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN ANTI
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI
; FILE REFERENCE: CHAOLIA
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-021-571-2

Query Match	100.0%;	Score 8853;	DB 14;	Length 1715;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1715;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAEGQNVIVKELKN	60	
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERECGQTPMLAAEGQNVIVKELKN	60	
Qy	61	GANCNLEDLDNWTALISASKEGHIHIVEELLKKSASLEHRDMGQWTALMWACYKGRDWW	120	
Db	61	GANCNLEDLDNWTALISASKEGHIHIVEELLKKSASLEHRDMGQWTALMWACYKGRDWW	120	
Qy	121	ELLLSHGANSVTLGQYSVYPIIWAAGRHADIVHLLIQNGAKVNCSDKYTTPLVWAAR	180	
Db	121	ELLLSHGANSVTLGQYSVYPIIWAAGRHADIVHLLIQNGAKVNCSDKYTTPLVWAAR	180	

121 ELLSHGANF SVTGLQY SVYPIIWAAGRGHADI V HLLQNGAKVNCSDKYGTTPLVWAAR 180
181 KGHLECVKHL LAMGADV DQEGANSMTALI I VAVGGY TQSVKEIL KRNPNVNLTKDGN TA 240
181 KGHLECVKHL LAMGADV DQEGANSMTALI I VAVGGY TQSVKEIL KRNPNVNLTKDGN TA 240
241 LMTASKEGHI EIVQDLLDAGTYNIPDRSGDTVLIGAVRGHVEI VRALLQKQVADIDIRG 300
241 LMTASKEGHI EIVQDLLDAGTYNIPDRSGDTVLIGAVRGHVEI VRALLQKQVADIDIRG 300
301 QDNKTALYNAVEKGNATMRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLDKGA 360
301 QDNKTALYNAVEKGNATMRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLDKGA 360
361 KVSADVKKGTPLHVAIRGRSRLAEILLRNPKDGRLLYRPNKAGETPNIDCSHQKSI L 420
361 KVSADVKKGTPLHVAIRGRSRLAEILLRNPKDGRLLYRPNKAGETPNIDCSHQKSI L 420
421 TQIFGARHLSPTEITDGMGLGYDLYSSALADILSEPTMQPPI CVGLYAQWGSFKSLKKL 480
421 TQIFGARHLSPTEITDGMGLGYDLYSSALADILSEPTMQPPI CVGLYAQWGSFKSLKKL 480
481 EDEKMTFAGQOTPELFOFQSWLIVFLTL LCGGLGVFAFPVDNTNLAIAISLSFLALIYIF 540
481 EDEKMTFAGQOTPELFOFQSWLIVFLTL LCGGLGVFAFPVDNTNLAIAISLSFLALIYIF 540
541 FIVYFGRRREGESWNWAWALSTRLARHIGYLELLFKLMFNVPPELPEQTTKALPVRF L 600
541 FIVYFGRRREGESWNWAWALSTRLARHIGYLELLFKLMFNVPPELPEQTTKALPVRF L 600
601 TDYNRLSSVGGETS LAEMIA TLSDACEREFGLATRLFRVFRTEESQKKWKKTCLPS 660
601 TDYNRLSSVGGETS LAEMIA TLSDACEREFGLATRLFRVFRTEESQKKWKKTCLPS 660
661 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVGGLAFVLCRTWQVLDSSL 720
661 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVGGLAFVLCRTWQVLDSSL 720
721 NSQKRLHSAASKLHLKSGFVKVCEVELMARMAKTIDSTONQTRLVWIIDGLDAC 780
721 NSQKRLHSAASKLHLKSGFVKVCEVELMARMAKTIDSTONQTRLVWIIDGLDAC 780
781 EQDKVLQMLDTRVRLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 840
781 EQDKVLQMLDTRVRLFSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 840
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
901 RRDYRROQRTITROMSFDLT KLLVTEWDFSDISPOTMRRLLINIVSVTGRLLRANQIT 960
901 RRDYRROQRTITROMSFDLT KLLVTEWDFSDISPOTMRRLLINIVSVTGRLLRANQIT 960
961 FNWDLASWNLTEQWYRPTSWLIIYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
961 FNWDLASWNLTEQWYRPTSWLIIYLEETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLPCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLPCTVNLDPKLEIITADVRAAREQINIGGLAY 1080
1081 PPLPLHGGPPRPPSGYQOPASVCSASFNGFPFGGVVSPOPHSYYSGLSGPOHPFNRA 1140
1081 PPLPLHGGPPRPPSGYQOPASVCSASFNGFPFGGVVSPOPHSYYSGLSGPOHPFNRA 1140
1141 AVPATGSSLLISSMTVDVVEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSCQNIIDELK 1200
1141 AVPATGSSLLISSMTVDVVEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSCQNIIDELK 1200
1201 KEMAMNFGDWHLFRSMVLEMRSVESQVVPDPRFLNENSGAPVPHGESARRSHTLEPLT 1260
1201 KEMAMNFGDWHLFRSMVLEMRSVESQVVPDPRFLNENSGAPVPHGESARRSHTLEPLT 1260

1261 ELSSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSOTRRTTPSLSSINQDSSIEISKLT 1320
1261 ELSSQTPYTLNPSFEELNTLGLDEGAPRHSNLSWQSOTRRTTPSLSSINQDSSIEISKLT 1320
1321 DKVQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIIGQSSSGGSIHSTLQERG 1380
1321 DKVQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYYIIGQSSSGGSIHSTLQERG 1380
1381 KEGELKQEGGRKSF LMKRGDVIDYSSSGVSTWEASPLDPI TEDEKSDQSGSKLLPGKKS 1440
1381 KEGELKQEGGRKSF LMKRGDVIDYSSSGVSTWEASPLDPI TEDEKSDQSGSKLLPGKKS 1440
1441 SERPSLFOQDLKLGGLRYQKLPDEDSGTRVOITPHCSQWITRKLKAKQRECAPS 1500
1441 SERPSLFOQDLKLGGLRYQKLPDEDSGTRVOITPHCSQWITRKLKAKQRECAPS 1500
1501 QEHSAPIRTFIKAKEYLSDALLDKXSDSGVRSNESSPNHSLHNEAADDSDQLEKANLI 1560
1501 QEHSAPIRTFIKAKEYLSDALLDKXSDSGVRSNESSPNHSLHNEAADDSDQLEKANLI 1560
1561 ELEDEGHSGRKGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPESWPACOKAYN 1620
1561 ELEDEGHSGRKGMPHSLGLQDPIIARMSICSEDKKSPSECSLIASSPESWPACOKAYN 1620
1621 LNRTPSTVTILNNNTAPTNRANQNFDEIGRETISOVLIRPGSPNPTAVQENLKSMAHK 1680
1621 LNRTPSTVTILNNNTAPTNRANQNFDEIGRETISOVLIRPGSPNPTAVQENLKSMAHK 1680
1681 RSQRSSYTRLSDASELHAASSESTGFEERSIL 1715
1681 RSQRSSYTRLSDASELHAASSESTGFEERSIL 1715

RESULT 2

US-10-117-229-6
; Sequence 6, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-6

Query Match 100.0%; Score 8853; DB 14; Length 1715;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSVLISQSVINYYEENIPALKALLEKCKVDNERNEGQTPLMLAAEQGNVEIVKELLKN 60
Db 1 MSVLISQSVINYYEENIPALKALLEKCKVDNERNEGQTPLMLAAEQGNVEIVKELLKN 60
Qy 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWAICYKGRDVV 120
Db 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWAICYKGRDVV 120
Qy 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADI V HLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADI V HLLQNGAKVNCSDKYGTTPLVWAAR 180
Qy 181 KGHLECVKHL LAMGADV DQEGANSMTALI I VAVGGY TQSVKEIL KRNPNVNLTKDGN TA 240
Db 181 KGHLECVKHL LAMGADV DQEGANSMTALI I VAVGGY TQSVKEIL KRNPNVNLTKDGN TA 240
Qy 241 LMTASKEGHI EIVQDLLDAGTYNIPDRSGDTVLIGAVRGHVEI VRALLQKQVADIDIRG 300


```
Db 241 LMIASKEGHEIIVQDLLDAGTYVNIIPDRSGDVTLIGAVRGHVEIVRALLQYADIDIRG 300
Qy 301 QDNKTALYWAKEGNATVRDILOCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 360
Db 301 QDNKTALYWAKEGNATVRDILOCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 360
Qy 361 KVSADVKKDGTPLHVAIRGRSRRRLAELLLRNPKOGLLYRPNKAGETPNYIDCSHQKSL 420
Db 361 KVSADVKKDGTPLHVAIRGRSRRRLAELLLRNPKOGLLYRPNKAGETPNYIDCSHQKSL 420
Qy 421 TQIFGARHLSSTETDGMGLVDLYSSALADILSEPTMQPPICVGLYAWGSGKFLKKL 480
Db 421 TQIFGARHLSSTETDGMGLVDLYSSALADILSEPTMQPPICVGLYAWGSGKFLKKL 480
Qy 481 EDEMKTFAGQQTPELPFQFSWLIIVFTLLCCGLGLVAFVVDTNLAISLSFALIIYIF 540
Db 481 EDEMKTFAGQQTPELPFQFSWLIIVFTLLCCGLGLVAFVVDTNLAISLSFALIIYIF 540
Qy 541 FIVYFGRREGESNNWAWALSTRARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
Db 541 FIVYFGRREGESNNWAWALSTRARHIGYLELLFKLMFVNPPPELPEQTTKALPVRFLF 600
Qy 601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEBSQKKWKTKCCLPS 660
Db 601 TDYNRLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFRTEBSQKKWKTKCCLPS 660
Qy 661 FVIFLFIAGCIIAGITLLAIIRVDPKHLTVNAIILSIASVVGGLAVFLNCRMTWQVLDL 720
Db 661 FVIFLFIAGCIIAGITLLAIIRVDPKHLTVNAIILSIASVVGGLAVFLNCRMTWQVLDL 720
Qy 721 NSQRKRLHSAASKLHLKSEGFMKVLKCEVELMARMKATIDSFQNTQRLVVIIDGLDAC 780
Db 721 NSQRKRLHSAASKLHLKSEGFMKVLKCEVELMARMKATIDSFQNTQRLVVIIDGLDAC 780
Qy 781 EQDKVQLQMDTVRVLFPSGPFIAIPASDPHIIIKAINONLSVLRDSNINGHDYMRIVH 840
Db 781 EQDKVQLQMDTVRVLFPSGPFIAIPASDPHIIIKAINONLSVLRDSNINGHDYMRIVH 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDRVSONSGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTGTQEDTDRVSONSGEMTKLGSKTALN 900
Qy 901 RRDYTRRQMQRTITRQMSFDLTKLLVTEDFWSDISQPMRRLNIVSVTGRLLRANQIT 960
Db 901 RRDYTRRQMQRTITRQMSFDLTKLLVTEDFWSDISQPMRRLNIVSVTGRLLRANQIT 960
Qy 961 FNRDLASWINLTQWPYRTSWLILYLBETEGLPDQMTLKTMYERI SKNIPTTKDVBPLL 1020
Db 961 FNRDLASWINLTQWPYRTSWLILYLBETEGLPDQMTLKTMYERI SKNIPTTKDVBPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPLVARVDKTPCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPLVARVDKTPCTVNLDPKLEIIADVRAAREQINIGGLAY 1080
Qy 1081 PPLPLHEGPPRPPSGYSPASVCSASFPNGPFGVSPQPHSSYVSGLSQPQHPFYNRA 1140
Db 1081 PPLPLHEGPPRPPSGYSPASVCSASFPNGPFGVSPQPHSSYVSGLSQPQHPFYNRA 1140
Qy 1141 AVPATGSSLLSSMTVDVVCSEKLQIEGLDQNMPOYCTTIKKANINGRVLSCQNBELK 1200
Db 1141 AVPATGSSLLSSMTVDVVCSEKLQIEGLDQNMPOYCTTIKKANINGRVLSCQNBELK 1200
Qy 1201 KEMANFGDHLFRSMVLEMSVSVQVPPDRFLNENSSAPVPHGESARSSHTPLPT 1260
Db 1201 KEMANFGDHLFRSMVLEMSVSVQVPPDRFLNENSSAPVPHGESARSSHTPLPT 1260
Qy 1261 ELSQTPYTLNFSPEELNTLGLDGAPRHSNLSQSOQTRTPPSLSLNSQDSSIEISKLT 1320
Db 1261 ELSQTPYTLNFSPEELNTLGLDGAPRHSNLSQSOQTRTPPSLSLNSQDSSIEISKLT 1320
Qy 1321 DKVQAEYRDVREYIAQMSQLEGGTGSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERG 1380
```

```
Db 1321 DKVQAEYRDVREYIAQMSQLEGGTGSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERG 1380
Qy 1381 KEGELKQEDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGRKSFIMKRGDVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLFQTDLKLKGGGLRYQKLPDSEDESGTGRVQITPHPCS KMIRTKRLKAKQRECA SP 1500
Db 1441 SERPSLFQTDLKLKGGGLRYQKLPDSEDESGTGRVQITPHPCS KMIRTKRLKAKQRECA SP 1500
Qy 1501 QEHSAPERTPIKAKYELSDALLDKDSSDGVSRNNESSPHSLHNEAADDSQLEKANLI 1560
Db 1501 QEHSAPERTPIKAKYELSDALLDKDSSDGVSRNNESSPHSLHNEAADDSQLEKANLI 1560
Qy 1561 ELEDEHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYN 1620
Db 1561 ELEDEHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYN 1620
Qy 1621 LNRTPSTVTLLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Db 1621 LNRTPSTVTLLNNNTAPTNRANQNFDEIEGIRETSQVILRPGPSPNPTAVQENLKSMAHK 1680
Qy 1681 RSQRSSYTRLSKDASELHAASSESTGFEERESIL 1715
Db 1681 RSQRSSYTRLSKDASELHAASSESTGFEERESIL 1715
```

RESULT 3

```
US-10-117-229-7
; Sequence 7, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1762
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-7
```

```
Query Match 97.6%; Score 8639.5; DB 14; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;
```

```
Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAAEQGNVEIVKELKKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAAEQGNVEIVKELKKN 60
Qy 61 GANCNLEDLNDWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDLNDWTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRTDVV 120
Qy 121 ELLLSHGANSVTCLOYSVYPIIWAAGRGHADI VHALLONGCAKUNCSDKYCTTFLVWAAR 180
Db 121 ELLLSHGANSVTCLOYSVYPIIWAAGRGHADI VHALLONGCAKUNCSDKYCTTFLVWAAR 180
Qy 181 KGHLECVKHLAMGADVDEQGANSMNTALIIVAKGGYTSQVKEILKRNPNVNLTDKGNNTA 240
Db 181 KGHLECVKHLAMGADVDEQGANSMNTALIIVAKGGYTSQVKEILKRNPNVNLTDKGNNTA 240
Qy 241 LMTASKEGHEIEIVQDLLDAGTYVNIIPDRSGDVTLIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMTASKEGHEIEIVQDLLDAGTYVNIIPDRSGDVTLIGAVRGHVEIVRALLQYADIDIRG 300
Qy 301 QDNKTALYWAKEGNATVRDILOCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 360
Db 301 QDNKTALYWAKEGNATVRDILOCNPDTEICTKDGTEPLIKATKORNIIEVVELLDKGA 360
```

361 KVSADVKKGDTPLHVAIRGRSRLABELLRNPKDGLLYRPNKAGTTPNIDCSHOKSIL 420
360 KVSADVKKGDTPLHVAIRGRSRLABELLRNPKDGLLYRPNKAGTTPNIDCSHOKSIL 419
421 TQIFGARHLSPTTGDGMGLYDYSALADIISEPTWQPPICVGLYAQMGSGKSFLLKKL 480
420 TQIFGARHLSPTTGDGMGLYDYSALADIISEPTWQPPICVGLYAQMGSGKSFLLKKL 479
481 EDEWKTFAQOQTPLFOFSLVFLTLILCGGLGLVFAFPVDTNLAIALISLFLALIIYIF 540
480 EDEWKTFAQOQTPLFOFSLVFLTLILCGGLGLVFAFPVDTNLAIALISLFLALIIYIF 539
541 FIVIFGGRREGSNNWAWALSTRLARHIGYIELLFLKLPVNPPELPEQTTKALPVRFPLF 600
540 FIVIFGGRREGSNNWAWALSTRLARHIGYIELLFLKLPVNPPELPEQTTKALPVRFPLF 599
601 TDYNELSSVGGETSLEAMIAATLSDACEREFGLATLFRVFRTEESQKKWKKTCCPLPS 660
600 TDYNELSSVGGETSLEAMIAATLSDACEREFGLATLFRVFRTEESQKKWKKTCCPLPS 659
661 FVIFLFTVGCIIAGITLLAIFRVPDKHLTVNAILISIASVVGGLAFVNLNCRTWQVLDLSLL 720
660 FVIFLFTVGCIIAGITLLAIFRVPDKHLTVNAILISIASVVGGLAFVNLNCRTWQVLDLSLL 719
721 NSQKRHLHSAASKLHKLKSGFMKVLKCEVELMARMAKTTIDSTQNTQRLVAVIIDGLDAC 780
720 NSQKRHLHSAASKLHKLKSGFMKVLKCEVELMARMAKTTIDSTQNTQRLVAVIIDGLDAC 779
781 BQDKVQMLDTRVFLFKSGPPIALFASDPHIIKAINQNLNSVLNDSNINGHYDMRNIVH 840
780 BQDKVQMLDTRVFLFKSGPPIALFASDPHIIKAINQNLNSVLNDSNINGHYDMRNIVH 839
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 899
901 RRDYTRRQWORTITRQMSFDLTKLLVTEDFWSDISQPTMRRLNIVSVTGRLLRANQIT 960
900 RRDYTRRQWORTITRQMSFDLTKLLVTEDFWSDISQPTMRRLNIVSVTGRLLRANQIT 959
961 FNWDLASWINLTCQWPYRTSWLLIYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPPLL 1020
960 FNWDLASWINLTCQWPYRTSWLLIYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPPLL 1019
1021 BIDGIRNFVPLSRTPVLVARDVKTFPLCTVNLDPKRLIADVPAAREQINIGGLAY 1080
1020 BIDGIRNFVPLSRTPVLVARDVKTFPLCTVNLDPKRLIADVPAAREQINIGGLAY 1079
1081 PPLPLHEGPPRPSPGVSOPASVCSASPNPGVSPQPHSSYVYSGLSGQPHFFYNNR - 1139
1080 PPLPLHEGPPRPSPGVSOPASVCSASPNPGVSPQPHSSYVYSGLSGQPHFFYNNR 1139
1140 -----AAPPATGSSILL 1151
1140 FPAPYLYTPRYVPGSQHLLISRSVKTSLPRDQNNGLPCDSGFNKRQOAAVPAATGSSILL 1199
1152 SMTVDVVECKLROTEGLDQNMPOYCTTIKANINGRVLSCNIDELKEMWANNFGDW 1211
1200 SMTVDVVECKLROTEGLDQNMPOYCTTIKANINGRVLSCNIDELKEMWANNFGDW 1259
1212 LFRSMVLEMRSVESQVVPDPDFLNNSSAPVPHGESARRSHTLPLTELSSQTPYTLN 1271
1260 LFRSMVLEMRSVESQVVPDPDFLNNSSAPVPHGESARRSHTLPLTELSSQTPYTLN 1319
1272 FSFEELNTLGLDEGARPHSNLSWQSQTRTPSLSLNSQDSSIEISKLTDKVQAEYRDAY 1331
1320 FSFEELNTLGLDEGARPHSNLSWQSQTRTPSLSLNSQDSSIEISKLTDKVQAEYRDAY 1379
1332 REYIAQMSOLEGSGTSTISGRSSPHSTYVYICQSSSGGSIHSTLQERKGEKELQEDGR 1391
1380 REYIAQMSOLEGSGTSTISGRSSPHSTYVYICQSSSGGSIHSTLQERKGEKELQEDGR 1439
1392 KGFMLKRGDVIDYSSSGSVTNEASPLDPIITEBDEKSDQSGSKLLPGKKSRRPSLFTQDL 1451

Db 1440 KSFLMKRGDVIDYSSSGSVTNEASPLDPIITEBDEKSDQSGSKLLPGKKSRRPSLFTQDL 1499
Qy 1452 KLKGGGLRYQKLPDDEDESGTRGVQIIPHCSCMTIRTKELKAK-ORECASPOEHSAPERT 1510
Db 1500 KLKGGGLRYQKLPDDEDESGTTEESDNTPL-LKDDDKKAEGKAERVKSP-EHSAEPERT 1557
Qy 1511 FTKAKEYLSDALLDKOSSDSGVRNNESSPNHSLHNEAADDQLEKANLIELEDEGHSGK 1570
Db 1558 FTKAKEYLSDALLDKOSSDSGVRNNESSPNHSLHNEAADDQLEKANLIELEDEGHSGK 1617
Qy 1571 RGMPHSLSGLOPDIIFARMISICSEDKSPSECSLIASSPESWPAQKAYNLNRPSTVTL 1630
Db 1618 RGMPHSLSGLOPDIIFARMISICSEDKSPSECSLIASSPESWPAQKAYNLNRPSTVTL 1677
Qy 1631 NNNTAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQSSSYTRL 1690
Db 1678 NNNTAPTNRANQNFDEIEGIRETSQVILRPGSPNPTAVQENLKSMAHKSQSSSYTRL 1737
Qy 1691 SKDASELHAASESTGFGEERESIL 1715
Db 1738 SKDASELHAASESTGFGEERESIL 1762

RESULT 4
US-10-021-571-4
; Sequence 4, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTIB
; FILE REFERENCE: CHA011A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-571-4

Query Match 93.4%; Score 8266; DB 14; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
Qy 1 MSVLIISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMAAEQGNLEIVKELKN 60
Db 1 MSVLIISQSVINYVEENIPALKALLEKCKDVDERNECGQTPLMAAEQGNLEIVKELKN 60
Qy 61 GANCNLEDLNTWALLISASKEGHHIVVELLKSGLSIEHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLNTWALLISASKEGHHIVVELLKSGLSIEHRDMGWTALMWACYKGRDVV 120
Qy 121 ELLLSHGANSVTLGQVSVYPIIWAAGRGHADIHLLLLQNGAKVNCSDKYGTTPPLVWAAR 180
Db 121 ELLLSHGANSVTLGQVSVYPIIWAAGRGHADIHLLLLQNGAKVNCSDKYGTTPPLVWAAR 180
Qy 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKTEILKNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKTEILKNPNVNLTKDGNNTA 240
Qy 241 LMTASKEGHHIEIVQDLDDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALIQYADIDIRG 300
Db 241 LMTASKEGHHIEIVQDLDDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALIQYADIDIRG 300
Qy 301 QDNKNTALYWAVEKGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 360
Db 301 QDNKNTALYWAVEKGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 360

```
Qy 361 KVSADVKKGDTPLHVAIRGRERRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQKSL 420
Db 361 KVSADVKKGDTPLHVAIRGRSKLAELLRNPKDGRLLYRPNKAGETPNIDCSHQKSL 420
Qy 421 TOIGARHLSHSTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKFLKLL 480
Db 421 TOIGARHLSHSTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKFLKLL 480
Qy 481 EDEMKTFAGQTEPLFQPSWLIIVFLTLCCGLGLVPAFPVDTNLAISLSFLALIYIF 540
Db 481 EDEMKTFAGQTEPLFQPSWLIIVFLTLCCGLGLVPAFPVDTNLAISLSFLALIYIF 540
Qy 541 FIVYFGRRGREGSNWAWALSTRLARHIGYLELLFKLMFVNPPPELPEQTTKALFVRFLF 600
Db 541 FIVYFGRRGREGSNWAWALSTRLARHIGYLELLFKLMFVNPPPELPEQTTKALFVRFLF 600
Qy 601 TDYNRSSVGETSLAEMIATLSDACEREFGLATRLFRVFRTRSESQKKKKTKTCCCLPS 660
Db 601 TDYNRSSVGETSLAEMIATLSDACEREFGLATRLFRVFRFKTEDTQKKKKTKTCCCLPS 660
Qy 720 FVIFLFIIGCIIAGITTLAIIPRVDPKHLTVNAIISIASVVGFLAVLNCRTMWQVLSLL 720
Db 720 FVIFLFIIGCIIAGITTLAIIPRVDPKHLTVNAIISIASVVGFLAVLNCRTMWQVLSLL 720
Qy 781 NSQRKRLHSAASKLHLKSEGFMKVYLKCEVELMARMAKTIDSFQONQRLVVIIDGLDAC 780
Db 781 NSQRKRLHSAASKLHLKSEGFMKVYLKCEVELMARMAKTIDSFQONQRLVVIIDGLDAC 780
Qy 840 EODKVLQMDTVRVLFSGKPIAIPASDPHIIKAINQNLNVLSDNSINGHDYMRNIVH 840
Db 840 EODKVLQMDTVRVLFSGKPIAIPASDPHIIKAINQNLNVLSDNSINGHDYMRNIVH 840
Qy 900 LPVFLNRSGLSNARKFLVTSATNGDITCSDTGTQEDTDRRVSONSLGEMTKLASKTALN 900
Db 900 LPVFLNRSGLSNARKFLVTSATNGDVPSCDITGIQEDADRRVSONSLGEMTKLASKTALN 900
Qy 960 RRDYTRRRQMORTITROMSFDTLLKLVTEDFWSDISPQTMERLLNIVSVTGRLLRANQIT 960
Db 960 RRDYTRRRQMORTITROMSFDTLLKLVTEDFWSDISPQTMERLLNIVSVTGRLLRANQIS 960
Qy 1020 FNVWDLASWINLTQWPYRTSWLIYLEETEGLDPMQTLKTMYERISKNIPTTKDVBPLL 1020
Db 961 FNVWDLASWINLTQWPYRTSWLIYLEETEGIDPMQTLKTIYERISKNIPTTKDVBPLL 1020
Qy 1080 EIDGDINRFVFLSSRTPVLVARDVKTFPLCTVNLDPKRLRIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDINRFVFLSSRTPVLVARDVKTFPLCTVNLDPKRLRIADVRAAREQISIGGLAY 1080
Qy 1140 PPLPLEHGPFRPPSGYSQPSVCSASFNGFPFGVVSQPHSSYYSGLSQPQHPFYNRA 1140
Db 1081 PPLPLEHGPFRAPSGYSQPPSVCSSTSFNGFPFAGVVSQPHSSYYSGMTGPQHPFYNRG 1140
Qy 1200 AVPATGSLSSMTVDVVCBKLQIIEGLDQNMPOYCTTIKKANINGRVLSCQNDIBELK 1200
Db 1141 SGPAFGPVLNLSNVDVAVCEKLQIIEGLDOSMLPOYCTTIKKANINGRVLSCQNDIBELK 1200
Qy 1260 KEMAMNFGDHLFRSVMLEMSVESQVVPEDPRFLNENSSAPVPHGESARSSHTPLPT 1260
Db 1201 KEMMNFQDHLFRSTVLEMNASHVVPEDPRFLSSSSGAPVPHGPARASHNELPHT 1260
Qy 1320 ELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSQSQTRTPPSLSNLSQDSSISEISKLT 1320
Db 1261 ELSSQTPYTLNFSFEELNTLGLDEGAPRHSNLSQSQTRTPPSLSNLSQDSSISEISKLT 1320
Qy 1380 DKVQAEYRDVREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGSGSIHSTLEBERG 1380
Db 1321 DKVQAEYRDVREYIAQMSQLEGGPGSTISGRSSPHSTYYIGQSSGSGSIHSTLEBERG 1380
Qy 1440 KEGELKQEDCKRKFMLKRGDVIDYSSGSVSNFASPLDPTITEDEKSDQSGSKLLPGKKS 1440
Db 1381 KQSEPKPDGGRKSFMLKRGDVIDYSSGSVSNFASPLDPTITEDEKSDQSGSKLLPGKKS 1440
```

```
Qy 1441 SERPSLFTQDLKLGGLRYQKLPDESDSTGRVQITPHCSKMIRTKRLKAKQREKASP 1500
Db 1441 SERPSLFTQDLKLGGLRYQKLPDESDSTGRVQITPHCSKMIRTKRLKAKQREKASP 1499
Qy 1501 QEHSAEPIRTFIKAKEVYLSDALDKDSSDSGVSRNESSPNHSLHNEAADDSOLEKANLI 1560
Db 1500 PEHSAEPIRTFIKAKEVYLSDALDKDSSDSGVSRNESSPNHSLHNEAADDSOLEKANLI 1559
Qy 1561 ELEDGHSKRGKMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYN 1620
Db 1560 ELEDGHSKRGKMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYN 1619
Qy 1621 LNRPSTVTLNNNTAPTNRANQNFDEIEGIRETQVILRPGPSPNPTAVONENLKSMAHK 1680
Db 1620 LNRPSTVTLNNNTAPTNRANQNFDEIEGIRETQVILRPGPSPNPTAVONENLKSMAHK 1679
Qy 1681 RSQSSVTRLSKDASELH-AASSESTGFEERESIL 1715
Db 1680 RSQSSVTRLSKDPPELHAAASSESTGFEERESIL 1715

RESULT 5
US-10-117-229-2
; Sequence 2, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidine220pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-2

Query Match 93.4%; Score 8266; DB 14; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

Qy 1 MSVLISQSVINYVEENIPALKALLEKCDVDNERECQTPMLAAEQGNVEIVKELIKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCDVDNERECQTPMLAAEQGNVEIVKELIKN 60
Qy 61 GANCNLEDLONWTALISASKEGHIHVEELLKSGASLEHRDMGWTALMWAICYKGRTDVV 120
Db 61 GANCNLEDLONWTALISASKEGHIHVEELLKCGVNLHRDMGWTALMWAICYKGRTDVV 120
Qy 121 ELLLSHGANSVTCLOYSVYPIIWAAGRHADIVHLLALONGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANSVTCLOYSVYPIIWAAGRHADIVHLLALONGAKVNCSDKYGTTPLVWAAR 180
Qy 181 KGHLECYKHLAMGADVQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240
Db 181 KGHLECYKHLAMGADVQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240
Qy 241 LMTASKEGHIIEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
Db 241 LMTASKEGHIIEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300
Qy 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGGETPLIKATKMRNIEVVELLLDKGA 360
Db 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGGETPLIKATKMRNIEVVELLLDKGA 360
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGRLLYRPNKAGETPNIDCSHQKSL 420
Db 361 KVSADVKKGDTPLHVAIRGRSKLAELLRNPKDGRLLYRPNKAGETPNIDCSHQKSL 420
Qy 421 TOIFGARHLSHSTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKFLKLL 480
Db 421 TOIFGARHLSHSTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKFLKLL 480
```

Db 421 TQIFGARHLSPTEGDMGLGYDLYSSALADILSEPTWQPICVGLYAQWGSKFLKKL 480
Qy 481 EDEMTFAGQOTBPLFQFQSWLIIVFLTLCCGGLGVPAFPVDNTLAIASISFLALIYIF 540
Db 481 EDEMTFAGQOIIBPLFQFQSWLIIVFLTLCCGGLGVPAFTVHNPGLIAVSLSFLALIYIF 540
Qy 541 FIVYFGRRGEGESWNWAWLSRLARHIGLYELLKLMFVNPELPPEQTTKALPVRELF 600
Db 541 FIVYFGRRGEGESWNWAWLSRLARHIGLYELLKLMFVNPELPPEQTTKALPVRELF 600
Qy 601 TDYRLSSVGGETSIAEMIAITLSDACERBEFGLATLRFVRFTESQGGKKWKTCCCLPS 660
Db 601 TDYRLSSVGGETSIAEMIAITLSDACERBEFGLATLRFVRFTEDTQGGKKWKTCCCLPS 660
Qy 661 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVGGLAFVNLNCRTWQVLDL 720
Db 661 FVIFLFTVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVGGLAFVNLNCRTWQVLDL 720
Qy 721 NSQKRLHSAASKLHKLSGEGMKVLKCEVELMARMAKTIDSTONQTRLVWIIDGLDAC 780
Db 721 NSQKRLHSAASKLHKLSGEGMKVLKCEVELMARMAKTIDSTONQTRLVWIIDGLDAC 780
Qy 781 EODKVLQMLDTRVLPFSKGFPIAFASDPHIIKAINONLSVLRDSNINGHDMYRNIHV 840
Db 781 EODKVLQMLDTRVLPFSKGFPIAFASDPHIIKAINONLSVLRDSNINGHDMYRNIHV 840
Qy 841 LPVFLNSRGLSNARKFVLTATNGDITCSDTTGTQEDTDRRVSONSIGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFVLTATNGDVPDSDTTGTQEDADRRVSONSIGEMTKLGSKTALN 900
Qy 901 RRDYRBRQORTITROMSFDLTLLVTEWDFSDISQPMWRLNINVSVGRLLRANQIT 960
Db 901 RRDYRBRQORTITROMSFDLTLLVTEWDFSDISQPMWRLNINVSVGRLLRANQIS 960
Qy 961 FNWDLASWNLTEQWYRISWLIYLEETEGPDMQTLKTMVERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWNLTEQWYRISWLIYLEETEGIPDMQTLKTIYERISKNIPTTKDVEPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPVILVARDVKFLPCTVNLDPKLEIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVILVARDVKFLPCTVNLDPKLEIADVRAAREQISIGGLAY 1080
Qy 1081 PPLPLHGGPRPPSGYQOPASVCSASFNGPFGVVPQPHSSYYSGLSGPQHPFNRA 1140
Db 1081 PPLPLHGGPRPPSGYQOPASVCSSTSFNGPFGVVPQPHSSYYSGMTGPQHPFNRG 1140
Qy 1141 AVPATGSLILSSMTVDVCEKLRQIEGLDQNMMPQYCTTIKKANINGRVLSCNIDELK 1200
Db 1141 SGPAQGPVLLNSLVNDAVCEKLRQIEGLDQMLPQYCTTIKKANINGRVLACNIDELK 1200
Qy 1201 KEMAMNFGDWHLPFRSMVLEKRSVQVVPDPRFLNENSNAPVPHGESARRSHTPLT 1260
Db 1201 KEMAMNFGDWHLPFRSTVLEKRNAESHVVPDPRFLSESSGPAHPGPARASHNELPHT 1260
Qy 1261 ELSSQTPYTLNFSFEELNTGLDGPAPRHSNLSQSQTRTPPSLSNSODSSSIEISKLT 1320
Db 1261 ELSSQTPYTLNFSFEELNTGLDGPAPRHSNLSQSQTRTPPSLSNSODSSSIEISKLT 1320
Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGTSGSTISGRSSPHSTYYIGQSSGSIHSTLEQERG 1380
Db 1321 DKVQAEYRDAYREYIAQMSQLEGGFGPTTISGRSSPHSTYYMGQSSGSIHSLNEQEGK 1380
Qy 1381 KEGELKQEDGRKSLFKMGKGDVIDYSSGVSNVNEASPLDPITEEDEKSDQSGKLLPGKKS 1440
Db 1381 KDSFPKPDGRKSLFKMGKGDVIDYSSGVSNVNDASPLDPITEEDEKSDQSGKLLPGKKS 1440
Qy 1441 SERPSLFTDULKGGGLRYOKLPSDESGTGRTVTPHCSKMRITKRLAKQRECAASP 1500
Db 1441 SERPSLFTDULKGGGLRYOKLPSDESGTGRTVTPHCSKMRITKRLAKQRECAASP 1500
Qy 1501 QEHSAEPRTTIFKAYEYLSADALLDKQSSDSGVRNNESSPNHSLHNEAADSQLEKANLI 1560
Db 1500 PEHSAEPRTTIFKAYEYLSADALLDKQSSDSGVRNNESSPNHSLHNEAADSQLEKANLI 1559

Qy 1561 ELEDGHSKRGMPHSLSGLOPITARMWISICSDKSPSECSLIASSPEESWPACOKAYN 1620
Db 1560 ELEDGHSKRGMPHSLSGLOPITARMWISICSDKSPSECSLIASSPEESWPACOKAYN 1619
Qy 1621 LNRTPSTVTLNNTAPTNRANQNFQIEGIRTSQVILFPGSPNPPTAYQVONENLKSMAHK 1680
Db 1620 LNRTPSTVTLNNTAPTNRANQNFQIEGIRTSQVILFPGSPNPPTAYQVONENLKSMAHK 1679
Qy 1681 RSQRSSYTLKSDASELH-AASSESTGFGEERESIL 1715
Db 1680 RSQRSSYTLKSDASELH-AASSESTGFGEERESIL 1715

RESULT 6

US-10-473-574-26
; Sequence 26, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
; APPLICANT: TANG, Y. Tom; YUE, Henry;
; APPLICANT: KHAN, Farrah A.; ISON, Craig H.;
; APPLICANT: BAUGHN, Mariah R.; WARREN, Bridget A.;
; APPLICANT: DUGGAN, Brendan M.; THANGAVELU, Kavitha;
; APPLICANT: HONCHELL, Cynthia D.; AZIMZAI, Yalda;
; APPLICANT: ELLIOTT, Vicki S.; BURFORD, Neil;
; APPLICANT: DING, Li; YUE, Huibin;
; APPLICANT: BECHA, Shanya; EMERLING, Brooke M.;
; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;
; APPLICANT: BANDMAN, Olga; LAL, Preeti G.;
; APPLICANT: LEE, Sally; GIETZEN, Kimberly J.;
; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
; APPLICANT: LEE, Ernestine A.; SWARNAKAR, Anita;
; APPLICANT: RING, Huijun Z.; JONES, Karen Anne
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0918 USN
; CURRENT APPLICATION NUMBER: US/10/473,574
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: PCT/US02/09288
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US 60/294,451
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/291,870
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/290,518
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/288,609
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/283,769
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/281,323
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/280,508
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2755454CD1
US-10-473-574-26

Query Match 93.4%; Score 8266; DB 16; Length 1715;
Best local similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;

Qy 1 MSVLISQSVINYVEENIPALKLEKCKVDNERNECGQTPMLAAEQGNVIVKELKN 60
Db 1 MSVLISQSVINYVEENIPALKLEKCKVDNERNECGQTPMLAAEQGNVIVKELKN 60

Qy 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLWAAAR 180
Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLWAAAR 180
Qy 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILLKRNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTOSVKEILLKRNPNVNLTKDGNNTA 240
Qy 241 LMIASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 241 LMIASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 301 QDNKTALYVAVEKGNATVRDILQCNPTETICTKGETPLIKATKRNIEVEELLKDKGA 360
Db 301 QDNKTALYVAVEKGNATVRDILQCNPTETICTKGETPLIKATKRNIEVEELLKDKGA 360
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRPNKGRLLYRPNKAGETPYNIDCSHQKSL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRLAELLRPNKGRLLYRPNKAGETPYNIDCSHQKSL 420
Qy 421 TQIFGARHLSPTETDGMGLDYSSALADILSPTMQPPICVGLIYAQWGSFKLLKYL 480
Db 421 TQIFGARHLSPTETDGMGLDYSSALADILSPTMQPPICVGLIYAQWGSFKLLKYL 480
Qy 481 EDEKMTAGQOTEPLFQFQSWLIVFLTLCCGLGLVFAFPVDTNLATISFLALYIF 540
Db 481 EDEKMTAGQOTEPLFQFQSWLIVFLTLCCGLGLVFAFPVDTNLATISFLALYIF 540
Qy 541 FIVYFGRGRGESWNAWALSTRLARHIGYLELLKLMFVNPPPELQTTKALPVRFPLF 600
Db 541 FIVYFGRGRGESWNAWALSTRLARHIGYLELLKLMFVNPPPELQTTKALPVRFPLF 600
Qy 601 TDYNRLSVGETSLAEMIATLSACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660
Db 601 TDYNRLSVGETSLAEMIATLSACEREFGLATRLFRVFRTEESQKKWKTCCLPS 660
Qy 661 FVIFLIVGCIAGITLLAIFRVPKHLVTNAILISTASVVGAFVLCNRTWQVLSLL 720
Db 661 FVIFLIVGCIAGITLLAIFRVPKHLVTNAILISTASVVGAFVLCNRTWQVLSLL 720
Qy 721 NSQKRLHSAASKLHLKSEGFVLMKCEVELMARMKTIIDFTQNTQRLVVIDGLDAC 780
Db 721 NSQKRLHSAASKLHLKSEGFVLMKCEVELMARMKTIIDFTQNTQRLVVIDGLDAC 780
Qy 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLNSVLDRSNGHDMYRNIVH 840
Db 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLNSVLDRSNGHDMYRNIVH 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTQEDTDRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTQEDTDRVSONSLGEMTKLGSKTALN 900
Qy 901 RRDYRRROMORTITROMSPDLTKLLVTEDEFSDISPTMRRLNIVSVTGLLRANQIT 960
Db 901 RRDYRRROMORTITROMSPDLTKLLVTEDEFSDISPTMRRLNIVSVTGLLRANQIT 960
Qy 961 FNWDLASWINLTQWPVRTSWLILYEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLTQWPVRTSWLILYEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020
Qy 1021 EIDGIRNFVFLSRPVLVARDVKVPLCTVNLDPKLRIBIADVRAAREQINIGGLAY 1080
Db 1021 EIDGIRNFVFLSRPVLVARDVKVPLCTVNLDPKLRIBIADVRAAREQINIGGLAY 1080
Qy 1081 PPLPLEGPPPPSPGSPASVCSASFGNGPFGVSPQPHSSVYSGLSGPQHPFNRA 1140
Db 1081 PPLPLEGPPPPSPGSPASVCSASFGNGPFGVSPQPHSSVYSGLSGPQHPFNRA 1140
Qy 1141 AVPATGSSLLISSMTVDVVCCKLRQIEGLDQNMMPQYCTTIKANINGRVLSQCNIDELK 1200

Db 1141 SGPAPOPVLNLSNDAVCEKLRQIEGLDQNMMPQYCTTIKANINGRVLSQCNIDELK 1200
Qy 1201 KEMANFGDMHLPFRSMVLEMRVSQVVPEDPRFLNENSSAPVPHGEGARSSHTPLPT 1260
Db 1201 KEMANFGDMHLPFRSMVLEMRVSQVVPEDPRFLNENSSAPVPHGEGARSSHTPLPT 1260
Qy 1261 ELSQTPYTLNFPSEELNTLGLDEGAPRHSNLSQSQOTRRTPSSLNSQDSSIBISKLT 1320
Db 1261 ELSQTPYTLNFPSEELNTLGLDEGAPRHSNLSQSQOTRRTPSSLNSQDSSIBISKLT 1320
Qy 1321 DKVOAERYDAYREYIAQMSQLEGGTSGSTISGRSPHSTYYIGQSSSGSSTHSTLEQERG 1380
Db 1321 DKVOAERYDAYREYIAQMSQLEGGTSGSTISGRSPHSTYYIGQSSSGSSTHSTLEQERG 1380
Qy 1381 KEGELKEDGRKSFMLKRGDVIDYSSGVSNTNEASPLDPIITEBEKSDQSGSKLLPGKKS 1440
Db 1381 KDSFPKDDGRKSFMLKRGDVIDYSSGVSNTNEASPLDPIITEBEKSDQSGSKLLPGKKS 1440
Qy 1441 SERPSLFTQDLKLGGLRYOKLPFDEDESGTGRVQITPHCSKMITRKLKAKQREKASP 1500
Db 1441 SERPSLFTQDLKLGGLRYOKLPFDEDESGTGRVQITPHCSKMITRKLKAKQREKASP 1500
Qy 1501 QEHSABEPIRPIKAKYLSDALLDKQSSDGVSRNENSSPHSLHNEAADDQSKANLI 1560
Db 1501 PEHSABEPIRPIKAKYLSDALLDKQSSDGVSRNENSSPHSLHNEAADDQSKANLI 1560
Qy 1561 ELEDEGHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYN 1620
Db 1561 ELEDDSHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYN 1620
Qy 1621 LNRTPTSVTLNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPTAVONENLKSMAHK 1680
Db 1621 LNRTPTSVTLNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPTAVONENLKSMAHK 1680
Qy 1681 RSQRSSYTRLSKODASELH-AASSESTGFGERRESIL 1715
Db 1681 RSQRSSYTRLSKODASELH-AASSESTGFGERRESIL 1715
RESULT 7
US-10-117-229-4
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220PC
; FILE REFERENCE: 9U 104 RI
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-4

Query Match 92.8%; Score 8212; DB 14; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;
Qy 1 MSVLISQSVINYYVEENIPALKALLEKCKVDNERNEGQOTPLMLAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINYYVEENIPALKALLEKCKVDNERNEGQOTPLMLAAEQGNVEIVKELKN 60
Qy 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Db 61 GANCNLELDNWTALISASKEGHIHVEELLKSGASLEHRDMGGWTALMWACYKGRDVTW 120
Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLWAAAR 180
Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNCSDKYGTTPLWAAAR 180
Qy 171 AVPATGSSLLISSMTVDVVCCKLRQIEGLDQNMMPQYCTTIKANINGRVLSQCNIDELK 1200

```
QY 181 KGHLECVKHLAMGADVDDQGANSM TAL I VAVKGYTQS VKELKRNPNVNLTKDGN TA 240
Db 180 KGHLECVKHLAMGADVDDQGANSM TAL I VAVKGYTQS VKELKRNPNVNLTKDGN TA 239
QY 241 LMIASKEGHIIEI VQDL DACTYNNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 300
Db 240 LMIASKEGHIIEI VQDL DACTYNNIPDRSGDTVLIGAVRGHVEIVRALLQKADIDIRG 299
QY 301 QDNKTALYWAVERKGNATWVRDILLOCNPDTEI CTYKDGETPLIKATKRNIEVBVELLDKGA 360
Db 300 QDNKTALYWAVERKGNATWVRDILLOCNPDTEI CTYKDGETPLIKATKRNIEVBVELLDKGA 359
QY 361 KVSADVKKGGTPTLHVAVRGSRRLAELLLRNPKDGLLYRPNKAGETPYNIDC SHOKS IL 420
Db 360 KVSADVKKGGTPTLHVAVRGSRRLAELLLRNPKDGLLYRPNKAGETPYNIDC SHOKS IL 419
QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGKSLKKL 480
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSFGKSLKKL 479
QY 481 EDEKTFAGQOTPLPFOFSLIIVFLTL LCGGLGLVPAFPVDNVLAI AISLFI ALIYIF 540
Db 480 EDEKTFAGQOI BPLFOFSLIIVFLTL LCGGLGLLFAFTVHPNLGIAVLSLFI ALIYIF 539
QY 541 FIVYFGRRREGSNNWAWLSRLARHIGVLELLFKLMFVNPPPELQOTTKALPVRF LF 600
Db 540 FIVYFGRRREGSNNWAWLSRLARHIGVLELLFKLMFVNPPPELQOTTKALPVRF LF 599
QY 601 TDYNRLLSVGGETS LAEMIA TLSDACEREFGLATRLFRVFRTEESQKKWKTKCLPS 660
Db 600 TDYNRLLSVGGETS LAEMIA TLSDACEREFGLATRLFRVFKETDQKKWKTKCLPS 659
QY 661 FVILFVLVGGIIAGITLLAI FRVDPKHLTVNAILLSIASVVG LAFVLNCR TWQVLD SLL 720
Db 660 FVILFVLVGGIIIGITLLAI FRVDPKHLTVNAILLSIASVVG LAFVLNCR TWQVLD SLL 719
QY 721 NSQRKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 780
Db 720 NSQRKRLHNAASKLHLKSGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 779
QY 781 EQDKVQLMDLTVRVLSKGFPIAIFASDPHII I KAINQNLSVLRDSNNGHDMYRNIVH 840
Db 780 EQDKVQLMDLTVRVLSKGFPIAIFASDPHII I KAINQNLSVLRDSNNGHDMYRNIVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVQNSIGEMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDVPDCSDTTGTQEDADRRVQNSIGEMTKLGSKTALN 899
QY 901 RRDYRRROMQRTITROMSFDLTKLLVTEDFWSDISPTQWRRLNIVSVTGRLLRANQIT 960
Db 900 RRDYRRROMQRTITROMSFDLTKLLVTEDFWSDISPTQWRRLNIVSVTGRLLRANQIS 959
QY 961 FNWDRLASWINL TEQMPYRTSWILYL EETEGLPDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 960 FNWDRLASWINL TEQMPYRTSWILYL EETEGIPDQMTLKTIVYERISKNIPTTKDVEPLL 1019
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEI IADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEI IADVRAAREQISIGGLAY 1079
QY 1081 PPLPLHEGPPRPSPGSQPVASVCSSAFNGFPFGVSPPOPHSYSGLSGQHPFVN-- 1138
Db 1080 PPLPLHEGPPRPSPGSQPVSCSSISFNGFPFGVSPPOPHSYSGMTGPOHPFNRP 1139
QY 1139 -----RAAVP 1143
Db 1140 FFAPLYLTPRYPGGSQHLISRPVSKTSLPRDQNGLEVIKEDAAEGLSSPTDSSRGSGP 1199
QY 1144 ATGSSLLLSMTVDVCEKLRQIEGLDQNMPOYCTTI KKANINGRVLSCNIDELKEM 1203
Db 1200 APGFVLLNSLVNDVACEKLRQIEGLDQSM LPOYCTTI KKANINGRVLACQNDIELKEM 1259
```

```
QY 1204 AMNFGDWHLFRSMVLEMRSVESOVVEDPRFLNENSSAPVPHGESARRSSHTLPLTELTS 1263
Db 1260 NMNFGDWHLFRSTVLEMRNAESHVVEDPRFLSESSSGPHAGEPARRASHNELPHTELTS 1319
QY 1264 SQTPTLNFSPFELNLTGLDEGAPRHSNLSWQSQTRRTTPSLSLNSQDSISIEISKLTDKV 1323
Db 1320 SQTPTLNFSPFELNLTGLDEGAPRHSNLSWQSQTRRTTPSLSLNSQDSISIEISKLTDKV 1379
QY 1324 QAEYRDAYREYTAOMSOLEGGTGSSTISGRSSPHSYTYIGOSSSGGSIHSTLFOERKEG 1383
Db 1380 QAEYRDAYREYTAOMSOLEGGPGSTTISGRSSPHSYTYMGSSSGGSIHNSLQEKGKDS 1439
QY 1384 ELKQEDGRKSFMLKMRGDVIDYSSGSVYTNEASPLDPI TEDEKSDOSGSKLLPGKKSSE 1443
Db 1440 EPKPDGRKSFMLKMRGDVIDYSSGSVYTNDASPLDPI TEDEKSDOSGSKLLPGKKSSE 1499
QY 1444 PSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHCSKMIRTKLKAQORECASPOEH 1503
Db 1500 SSLFQTDLKLKGGSLRYQKLPSEDESGTGEESDNTP-LLKDDKDKRAEGKVERVPKSP EH 1558
QY 1504 SAEPIRTFKAEYLSDALDKKSDSGVRSNESPNSHSLHNEAADDSOLEKANLIELE 1563
Db 1559 SAEPIRTFKAEYLSDALDKKSDSGVRSSESPNSHSLHNEVADDSOLEKANLIELE 1618
QY 1564 DEHSGKRGMPHSLSLGLQDPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNR 1623
Db 1619 DSHSGKRGIPHSLSLGLQDPIIARMSICSEDKSPSECSLIASSPEENWPACOKAYNLNR 1678
QY 1624 TPSTVTNNAPTNRANQNFDEIGRETQSIVLAPGPSNPPTAVONENLKSMAHKRSQ 1683
Db 1679 TPSTVTNNAPSAPNRANQNFDEMEGIRETSQVILAPSSSPNPTTIQENLKSMTHKRSQ 1738
QY 1684 RSSYTRLSKDASELH-AASSESTGFGEERESIL 1715
Db 1739 RSSYTRLSKDPPHELHAAASSESTGFGEERESIL 1771

RESULT 8
US-10-117-229-11
; Sequence 11, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-117-229-11

Query Match 92.8%; Score 8212; DB 14; Length 1771;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 1590; Conservative 53; Mismatches 70; Indels 60; Gaps 4;

QY 1 MSVLIISQSVINYVEENIPALKALLEKCKDVBERNECGQTPLMLAAEQGNVEIVKELLKN 60
Db 1 MSVLIISQSVINYVEENIPALKALLEKCKDVBERNECGQTPLMLAAEQGNLEIVKELIKN 60
QY 61 GANCNLELDNNTALISASKEGHIHIVEELLKSGASLEHRDMCGGTALMWACYKGRTDVV 120
Db 61 GANCNLELDNNTALISASKEGHVHIVEELLKCGVNLHHRDMGMGTALMWACYKGRTDVV 120
QY 121 ELLLSHGANPSVTGLQYSVYPIIWAAGRGHADI VHLHLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANPSVTGL-YSVYPIIWAAGRGHADI VHLHLLQNGAKVNCSDKYGTTPLVWAAR 179
QY 181 KGHLECVKHLAMGADVDDQGANSM TAL I VAVKGYTQS VKELKRNPNVNLTKDGN TA 240
```


Db 180 KGHLECVKHLAMGADVDOEGANSMTALI VAVKGYTQSVKEILKRNPNVNLTDKGNTA 239
Qy 241 LMIASKEGHEIIVQDLIDAGTYVNI PDRSGDTVLICAVRGHVEIVRALLQYADIDIRG 300
Db 240 LMIASKEGHEIIVQDLIDAGTYVNI PDRSGDTVLICAVRGHVEIVRALLQYADIDIRG 299
Qy 301 QNKATLYAWVEKGNATVRDILQCNPDTEICTDGETPLIKATQORNIIEVVELLDKGA 360
Db 300 QNKATLYAWVEKGNATVRDILQCNPDTEICTDGETPLIKATQORNIIEVVELLDKGA 359
Qy 361 KVSADVKKGDTPHVAIRGRSPRLAELLRNPKDGLLYRKNKAGETPNYDCHQKSL 420
Db 360 KVSADVKKGDTPHVAIRGRSPRLAELLRNPKDGLLYRKNKAGETPNYDCHQKSL 419
Qy 421 TQIFGARHLSPTETDGMGLDYSSALADILSEPTMOPPICVGLIYAQWGSKSLKXL 480
Db 420 TQIFGARHLSPTETDGMGLDYSSALADILSEPTMOPPICVGLIYAQWGSKSLKXL 479
Qy 481 EDEMTFAGQOETPLFQPSWLIIVFTLLCGGLGVAPPVDTNIAIASISFLALIYIF 540
Db 480 EDEMTFAGQOETPLFQPSWLIIVFTLLCGGLGVAPPVDTNIAIASISFLALIYIF 539
Qy 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLFKLMFVNPPPELPTQTTKALPVRELF 600
Db 540 FIVYFGRREGESNNWAWLSTRLARHIGYLELLFKLMFVNPPPELPTQTTKALPVRELF 599
Qy 601 TDYNRLSSVGGETSLAEMIATLSACEREFGLFATRLFRVFRTERESQKKKKWKTCCPLPS 660
Db 600 TDYNRLSSVGGETSLAEMIATLSACEREFGLFATRLFRVFRTERESQKKKKWKTCCPLPS 659
Qy 661 FVIFLFIIGCIIGITLAIIRVDPKHLTVNAVLISIASVVGGLAPVLCRTWQVLDL 720
Db 660 FVIFLFIIGCIIGITLAIIRVDPKHLTVNAVLISIASVVGGLAPVLCRTWQVLDL 719
Qy 721 NSQKRLHSAASKLHLKSEGFMYKLVCEVELMARMKTIIDSTFQNTQRLVVIDGLDAC 780
Db 720 NSQKRLHSAASKLHLKSEGFMYKLVCEVELMARMKTIIDSTFQNTQRLVVIDGLDAC 779
Qy 781 EODKVLQMDTVRVLFSGPFIAPASDPHIIKAINQNLNVLSDNSINGHDYRNIVH 840
Db 780 EODKVLQMDTVRVLFSGPFIAPASDPHIIKAINQNLNVLSDNSINGHDYRNIVH 839
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVVSQNSLGEMTKLASKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVVSQNSLGEMTKLASKTALN 899
Qy 901 RRDYRRRQMTITRQMSFDTKLLVTEWFDSDISQPMRLNINIVSVTGRLLRANQIT 960
Db 900 RRDYRRRQMTITRQMSFDTKLLVTEWFDSDISQPMRLNINIVSVTGRLLRANQIS 959
Qy 961 FNMRLASWINLTQWMPYRTSWLILYLBETEGLPDQMTLKTMYERISQNIPTTKDVEPLL 1020
Db 960 FNMRLASWINLTQWMPYRTSWLILYLBETEGLPDQMTLKTMYERISQNIPTTKDVEPLL 1019
Qy 1021 EIDGDIRNFVFLSSRTPLVARDVKTPLCTVNLDPKRLIADVRAAREQINIGGLAY 1080
Db 1020 EIDGDIRNFVFLSSRTPLVARDVKTPLCTVNLDPKRLIADVRAAREQINIGGLAY 1079
Qy 1081 PPLPLHEGPPRPSPGYOPASVCSASFNPGFPGWSPQPHSSYSGLSGQPHFFYN-- 1138
Db 1080 PPLPLHEGPPRPSPGYOPASVCSASFNPGFPGWSPQPHSSYSGLSGQPHFFYNRP 1139
Qy 1139 -----RAVP 1143
Db 1140 FFAPLYTPRYPGGSQHLISRPVKTSLPRDQNGLEVIKEDAAEGSLSPDSSRGSGP 1199
Qy 1144 ATGSLYLLSMTVDVVCBKLQIEGLDQNMMPQVCTTIKKNINGRVLSCQNIIDELKEM 1203
Db 1200 ACPVLLNSLVNDAVCEKQLQIEGLDQSMPLQVCTTIKKNINGRVLSCQNIIDELKEM 1259
Qy 1204 AMNFGDWHLFRSMVLENRSVESQVVPDPREFLNENSSAPVPHGSARRSHTPLTFLS 1263
Db 1260 NMNFGDWHLFRSTVLEMRNASHVVPDPREFLSESSSGPA PHGHPARRASHNELPHTLS 1319

Qy 1264 SOTPYTLNFSPEELNTLGLDEGAPHSNLSWQSOTRRTPSLSSLSNSQSSSIKSLTDKV 1323
Db 1320 SOTPYTLNFSPEELNTLGLDEGAPHSNLSWQSOTRRTPSLSSLSNSQSSSIKSLTDKV 1379
Qy 1324 QAEYRDAYREYIAQWQLEGGTSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERKRG 1383
Db 1380 QAEYRDAYREYIAQWQLEGGTSGSTISGRSSPHSTYYIGOSSGGSIHSTLEQERKRG 1439
Qy 1384 ELKQEDGKSKFLMKRGVDIYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER 1443
Db 1440 EPKDDGGRKSKFLMKRGVDIYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER 1499
Qy 1444 PSLFOTDLKLGKGLRYOKLPDSDESGTGRVQITPHCSKMIRTKRLKAKQREKASPOEH 1503
Db 1500 SSLFOTDLKLGKGLRYOKLPDSDESGTGRVQITPHCSKMIRTKRLKAKQREKASPOEH 1558
Qy 1504 SAEPIRTFIKAKYLSLSDALLDKDSSDGVSRNNESSPNHSLHNEAADDQSOLEKANLIELE 1563
Db 1559 SAEPIRTFIKAKYLSLSDALLDKDSSDGVSRNNESSPNHSLHNEAADDQSOLEKANLIELE 1618
Qy 1564 DEHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNR 1623
Db 1619 DSHSGKRGMPHSLSGLODPIIARMSICSEDKKSPSECSLIASSPEESWPACQKAYNLNR 1678
Qy 1624 TPTVTTLNNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPNTAVONENLKSMAHKSQ 1683
Db 1679 TPTVTTLNNNTAPTNRANQNFDEIGRETQSIVILRPGSPNPNTAVONENLKSMAHKSQ 1738
Qy 1684 RSYSTRLSKQASELH-AASSESTGFEERESIL 1715
Db 1739 RSYSTRLSKQASELH-AASSESTGFEERESIL 1771

RESULT 9

US-10-117-229-3
; Sequence 3, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION: Technologies, Inc.
; APPLICANT: Origene Technologies, Inc.
; FILE OF INVENTION: Human Kidins220Pc
; TITLE OF INVENTION: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; TYPE: PRT
; LENGTH: 1184
; ORGANISM: Homo sapiens
US-10-117-229-3

Query Match 64.7%; Score 5730.5; DB 14; Length 1184;

Best Local Similarity 96.2%; Pred. No. 0;

Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINYYVEENIPALKALLEKQVDNERECGQTPMLAAEOGNVIEVKELKN 60
Db 1 MSVLISQSVINYYVEENIPALKALLEKQVDNERECGQTPMLAAEOGNVIEVKELKN 60
Qy 61 GANCNLEDLQNTALISASKEGHIHVEELLKSGASLEHRDMGGMGTALMWACYKGRDVV 120
Db 61 GANCNLEDLQNTALISASKEGHIHVEELLKSGASLEHRDMGGMGTALMWACYKGRDVV 120
Qy 121 ELLLSHGANSVTLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLLSHGANSVTLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 179
Qy 181 KGHLECVKHLAMGADVDOEGANSMTALI VAVKGYTQSVKEILKRNPNVNLTDKGNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALI VAVKGYTQSVKEILKRNPNVNLTDKGNTA 239
Qy 241 LMIASKEGHEIIVQDLIDAGTYVNI PDRSGDTVLICAVRGHVEIVRALLQYADIDIRG 300

240 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
301 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKOGETPLIKATKORNIIEVVELLDKGA 360
300 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKOGETPLIKATKORNIIEVVELLDKGA 359
361 KVSANDKGGDTPLHVAIRGSRRLAEALLRNPKDGRLLYRPNKAGETPNYDNCSHOKSIL 420
360 KVSANDKGGDTPLHVAIRGSRRLAEALLRNPKDGRLLYRPNKAGETPNYDNCSHOKSIL 419
421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 480
420 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 479
481 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVPAFPVDTNLAIAISLSFLALIYIF 540
480 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVPAFPVDTNLAIAISLSFLALIYIF 539
541 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 600
540 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 599
601 TDYNRSLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCPLS 660
600 TDYNRSLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCPLS 659
661 FVIFLFTVGCIIAGITLLAIFRVDPKHLTNAILISIASVVGGLAFVNLNCTWQVLDLSLL 720
660 FVIFLFTVGCIIAGITLLAIFRVDPKHLTNAILISIASVVGGLAFVNLNCTWQVLDLSLL 719
721 NSQKRLHSAASKLHKLKSGFMMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 780
720 NSQKRLHSAASKLHKLKSGFMMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 779
781 EQDKVQLMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINNGHYDMRNVH 840
780 EQDKVQLMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINNGHYDMRNVH 839
841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSGEMTKLGSKTALN 900
840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSGEMTKLGSKTALN 899
901 RRDYRRRQMQRTITROMSFDLTCLAVTEDWFSDISPQTMERLLNIVSVTGRLLRANQIT 960
900 RRDYRRRQMQRTITROMSFDLTCLAVTEDWFSDISPQTMERLLNIVSVTGRLLRANQIT 959
961 FNWDRLASWINLTSQWPFRTSWLLIYLEETEGLPDQMTLKTMYERI SKNIPTTKDVEPLL 1020
960 FNWDRLASWINLTSQWPFRTSWLLIYLEETEGIPDQMTLKTMYERI SKNIPTTKDVEPLL 1019
1021 EIDGDIRNFVFLSSRTPVLVARVDTPLCTVNLDPKLEIITADVRAAEQINIGGLAY 1080
1020 EIDGDIRNFVFLSSRTPVLVARVDTPLCTVNLDPKLEIITADVRAAEQISIGGLAY 1079
1081 PPLPLHGGPPRPPSGYSGOPASVCSASFNGPFGGVVSPPHSSYYSGLSGPOHPFYNR 1139
1080 PPLPLHGGPPRPPSGYSGOPASVCSASFNGPFGGVVSPPHSSYYSGMTGPQHPFYNR 1138

RESULT 10

US-10-117-229-9
; Sequence 9, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117, 229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9

; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-9

Query Match 64.7%; Score 5730.5; DB 14; Length 1184;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1096; Conservative 24; Mismatches 18; Indels 1; Gaps 1;

QY 1 MSVLISQSIVNYVEENIPALKALIEKCKXVDVNERECGOTPLMLAAEQGNVETVKELKN 60
DB 1 MSVLISQSIVNYVEENIPALKALIEKCKXVDVNERECGOTPLMLAAEQGNLEIVKELKN 60
QY 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGASLEHRDMGCGTALMAWACYKGRDVV 120
DB 61 GANCNLEDLNNWTALISASKEGHIHIVEELLKSGVLEHRDMGCGTALMAWACYKGRDVV 120
QY 121 ELLSHGANPSVTGLQYVYPIIWAAGRGHADIIVHLLQNGAKVNSDKYGTTPVWAAR 180
DB 121 ELLSHGANPSVTGL-QYVYPIIWAAGRGHADIIVHLLQNGAKVNSDKYGTTPVWAAR 179
QY 181 KGHLECVKHLLANGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKOGNTA 240
DB 180 KGHLECVKHLLANGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKOGNTA 239
QY 241 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
DB 240 LMIASKEGHEIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKOGETPLIKATKORNIIEVVELLDKGA 360
DB 300 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKOGETPLIKATKORNIIEVVELLDKGA 359
QY 361 KVSANDKGGDTPLHVAIRGSRRLAEALLRNPKDGRLLYRPNKAGETPNYDNCSHOKSIL 420
DB 360 KVSANDKGGDTPLHVAIRGSRRLAEALLRNPKDGRLLYRPNKAGETPNYDNCSHOKSIL 419
QY 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 480
DB 420 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFKLLKCL 479
QY 481 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVPAFPVDTNLAIAISLSFLALIYIF 540
DB 480 EDEMKTFAGQOETPLFQFWSLIVFLTLCCGGLGLVPAFPVDTNLAIAISLSFLALIYIF 539
QY 541 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 600
DB 540 FIVYFGRREGESWNWAWLSTRLARHIGVLELLPKLMFVNPELPEQTTKALPVRFLLF 599
QY 601 TDYNRSLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCPLS 660
DB 600 TDYNRSLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCCPLS 659
QY 661 FVIFLFTVGCIIAGITLLAIFRVDPKHLTNAILISIASVVGGLAFVNLNCTWQVLDLSLL 720
DB 660 FVIFLFTVGCIIAGITLLAIFRVDPKHLTNAILISIASVVGGLAFVNLNCTWQVLDLSLL 719
QY 721 NSQKRLHSAASKLHKLKSGFMMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 780
DB 720 NSQKRLHSAASKLHKLKSGFMMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC 779
QY 781 EQDKVQLMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINNGHYDMRNVH 840
DB 780 EQDKVQLMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINNGHYDMRNVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSGEMTKLGSKTALN 900
DB 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSGEMTKLGSKTALN 899
QY 901 RRDYRRRQMQRTITROMSFDLTCLAVTEDWFSDISPQTMERLLNIVSVTGRLLRANQIT 960
DB 900 RRDYRRRQMQRTITROMSFDLTCLAVTEDWFSDISPQTMERLLNIVSVTGRLLRANQIS 959

Qy 961 FNRDLASWINLQWVYRTSWLILYLEETEGLPDQMTLKTMYERISKNIPPTTKDVEPLL 1020
 Db 960 FNRDLASWINLQWVYRTSWLILYLEETEGIPDQMTLKTMYERISKNIPPTTKDVEPLL 1019
 Qy 1021 EIDGDIRNFVFLSRTPLVARDVKTPLPCTVNLDPKRLIADVRAAREQINIGGLAY 1080
 Db 1020 EIDGDIRNFVFLSRTPLVARDVKTPLPCTVNLDPKRLIADVRAAREQISIGGLAY 1079
 Qy 1081 PPLPLHGGPPPPSGYSQPASVCSASFNPGVSPQPHSSYSYGLSGPQHPFYNR 1139
 Db 1080 PPLPLHGGPPPSGYSQPPSCSTSNFAGGVSPQPHSSYSYSGMTQHPFYNR 1138

RESULT 11
 US-10-106-698-6378
 ; Sequence 6378, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005P1
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 6378
 ; LENGTH: 705
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (244)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (337)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-10-106-698-6378

Query Match 33.5%; Score 2961.5; DB 14; Length 705;
 Best Local Similarity 95.5%; Pred. No. 1e-195;
 Matches 567; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

Qy 62 ANCNLEDLDNWTALISASKEGHIHVEELLKSGASLEHRDGGWTALMWACYKGRDVE 121
 Db 28 SNCNLEDLDNWTALISASKEGHVHVEELLKCGVNLHRDGGWTALMWACYKGRDVE 87
 Qy 122 LLLSHGANPSVTGLQYSYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPPLVWAARK 181
 Db 88 LLLSHGANPSVTGL-QYSYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPPLVWAARK 146
 Qy 182 GHLECVKHLAMGADVDEGANSMTALIVAKGGVTSVKELKRNPNVNLTKDGNL 241
 Db 147 GHLECVKHLAMGADVDEGANSMTALIVAKGGVTSVKELKRNPNVNLTKDGNL 206
 Qy 242 MIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 301
 Db 207 MIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 266
 Qy 302 DNKTALYWAKEGNATVNRDILQCNPDTEICTDGETPLIKATKORNIIEVVELLDKGA 361
 Db 267 DNKTALYWAKEGNATVNRDILQCNPDTEICTDGETPLIKATKORNIIEVVELLDKGA 326
 Qy 362 VSAVDKGGDTPHVAIRGRSRRLAELLRNPKGRLLYRPNKAGETPNIDCSHQSILT 421
 Db 327 VSAVDKGGDTPHVAIRGRSRRLAELLRNPKGRLLYRPNKAGETPNIDCSHQSILT 386
 Qy 422 QIFGARHLSPTETDGMGLYSSALADILSEPTMQPPICVGLYAQWGSCKSFLKKLE 481

Db 387 QIFGARHLSPTETDGMGLYSSALADILSEPTMQPPICVGLYAQWGSCKSFLKKLE 446
 Qy 482 DEMKTFAGQTEPLPQFQSWLIVFLTLCCGGLGVFAFPVDTNLAIALISFLALYIFF 541
 Db 447 DEMKTFAGQTEPLPQFQSWLIVFLTLCCGGLGLLFAFVHPNGLGIAVSLSLALLYIFF 506
 Qy 542 IVIYFGRRREGESNNWAWALSTRLARHIGYLELFLKLMFVNPPPEQTITKALPVRFLLT 601
 Db 507 IVIYFGRRREGESNNWAWLSTRLARHIGYLELFLKLMFVNPPPEQTITKALPVRFLLT 566
 Qy 602 DYNRLSSVGGTSLAEMIATLSDCEREFGLPFLATRLFRVPTBSSQGGKKWKKT 655
 Db 567 DYNRLSSVGGTSLAEMIATLSDCEREFGLPFLATRLFRVPTBSSQGGKKKKNS 620

RESULT 12
 US-09-835-788A-17
 ; Sequence 17, Application US/09835788A
 ; Patent No. US20020077458A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ni et al.
 ; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
 ; FILE REFERENCE: PT018P1
 ; CURRENT APPLICATION NUMBER: US/09/835,788A
 ; CURRENT FILING DATE: 2001-04-17
 ; PRIOR APPLICATION NUMBER: PCT/US00/28666
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/159,585
 ; PRIOR FILING DATE: 1999-10-18
 ; PRIOR APPLICATION NUMBER: 60/167,246
 ; PRIOR FILING DATE: 1999-11-24
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Patent in ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 551
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-835-788A-17

Query Match 31.4%; Score 2782.5; DB 9; Length 551;
 Best Local Similarity 96.7%; Pred. No. 1.6e-183;
 Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Qy 102 MGGTALMWACYKGRDQVVELLSHGANPSVTGLQYSYPIIWAAGRHADIVHLLQNG 161
 Db 1 MGGTALMWACYKGRDQVVELLSHGANPSVTGL-QYSYPIIWAAGRHADIVHLLQNG 59
 Qy 162 AKVNCSDKYGTTPPLVWAARKGHLECVKHLAMGADVDEGANSMTALIVAKGGVTSVK 221
 Db 60 AKVNCSDKYGTTPPLVWAARKGHLECVKHLAMGADVDEGANSMTALIVAKGGVTSVK 119
 Qy 222 EILKRNPNVNLTKDGNLMIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRG 281
 Db 120 EILKRNPNVNLTKDGNLMIASKEGHIETVQDLDAGTYVNPDRSGDTVLIGAVRG 179
 Qy 282 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATVNRDILQCNPDTEICTDGETPLI 341
 Db 180 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATVNRDILQCNPDTEICTDGETPLI 239
 Qy 342 KATKORNIIEVVELLDKGAQVSAVDKGGDTPHVAIRGRSRRLAELLRNPKGRLLYR 401
 Db 240 KATKORNIIEVVELLDKGAQVSAVDKGGDTPHVAIRGRSRRLAELLRNPKGRLLYR 299
 Qy 402 NKAGETPNIDCSHQSILTQIFGARHLSPTETDGMGLYSSALADILSEPTMQPPI 461
 Db 300 NKAGETPNIDCSHQSILTQIFGARHLSPTETDGMGLYSSALADILSEPTMQPPI 359
 Qy 462 CVGLYAQWGSCKSFLKKLEDEMKTTFAGQTEPLPQFQSWLIVFLTLCCGGLGVFAFPV 521
 Db 360 CVGLYAQWGSCKSFLKKLEDEMKTTFAGQTEPLPQFQSWLIVFLTLCCGGLGVFAFPV 419

Qy 522 DTNLAIAISLFLALYIFPIVYFGRRGESNNWAWALSTRLARHIGYLELFLKMFV 581
Db 420 HPNLGIAVLSLFLALYIFPIVYFGRRGESNNWAWLSTRLARHIGYLELFLKMFV 479
Qy 582 NPPELPQTTKALPVRLFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 641
Db 480 NPPELPQTTKALPVRLFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 539
Qy 642 RTEESQKKKKW 653
Db 540 KTEDTQKKKK 551

RESULT 13

US-10-175-042-17
; Sequence 17, Application US/10175042
; Publication No. US2003018170A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides,
; TITLE OF INVENTION: Polypeptides, and
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/10/175,042
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 09/835,788
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-175-042-17

Query Match 31.4%; Score 2782.5; DB 14; Length 551;
Best Local Similarity 96.7%; Pred. No. 1.6e-183;
Matches 534; Conservative 8; Mismatches 9; Indels 1; Gaps 1;
Qy 102 MGGWTALMWAACYGRTDVVELLSHGANGPSVTGL-YSVYPIIWAAGRGHADI VHL L L QNG 59
Db 1 MGGWTALMWAACYGRTDVVELLSHGANGPSVTGL-YSVYPIIWAAGRGHADI VHL L L QNG 59
Qy 162 AKVNCSDKYCTTPLVWAARKHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVK 221
Db 60 AKVNCSDKYCTTPLVWAARKHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVK 119
Qy 222 EILKRNPNVNLTKDGNLTALMIASKEGHIIVQDLDAGTYVNI PDRSGDTVLIGAVRG 281
Db 120 EILKRNPNVNLTKDGNLTALMIASKEGHIIVQDLDAGTYVNI PDRSGDTVLIGAVRG 179
Qy 282 HVEIVRALLQYADIDIRGOONKTALYWEKGNATVVRDILQCNPDTEICTKDGTEPLI 341
Db 180 HVEIVRALLQYADIDIRGOONKTALYWEKGNATVVRDILQCNPDTEICTKDGTEPLI 239
Qy 342 KATKARNIEVVLELLDKGKVSADKKGDTPLHVAIRGRSRLAEILLRNPDKGRLLYRP 401
Db 240 KATKARNIEVVLELLDKGKVSADKKGDTPLHVAIRGRSRLAEILLRNPDKGRLLYRP 299
Qy 402 NKAGETPNYDCSHQSKSILTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI 461
Db 300 NKAGETPNYDCSHQSKSILTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI 359
Qy 462 CVGLIYAQWGSKGKSFLLKKLEDEMKTFAGQQTPELPQFSWLI VFL TLL L CGGLGLVFAFPV 521
Db 360 CVGLIYAQWGSKGKSFLLKKLEDEMKTFAGQQTPELPQFSWLI VFL TLL L CGGLGLVFAFTV 419

Qy 522 DTNLAIAISLFLALYIFPIVYFGRRGESNNWAWALSTRLARHIGYLELFLKMFV 581
Db 420 HPNLGIAVLSLFLALYIFPIVYFGRRGESNNWAWLSTRLARHIGYLELFLKMFV 479
Qy 582 NPPELPQTTKALPVRLFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 641
Db 480 NPPELPQTTKALPVRLFTDYNRLSSVGGETSLAEMIATLSDACERFGLATRLPRVF 539
Qy 642 RTEESQKKKKW 653
Db 540 KTEDTQKKKK 551

RESULT 14

US-10-021-571-8
; Sequence 8, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTIB
; FILE REFERENCE: CHAO11A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Drosophila
US-10-021-571-8

Query Match 26.8%; Score 2376.5; DB 14; Length 1498;
Best Local Similarity 37.6%; Pred. No. 1e-154;
Matches 552; Conservative 272; Mismatches 527; Indels 117; Gaps 36;
Qy 1 MSVLISQSVINYYEENIPALKALLE-KCKVDVERNECGQTPPLMLAABQGNVIVKELLK 59
Db 23 MGSGLHRALLQYIDNNDISGLRAILDSRLHITIDDRDENATTVLMVAVAGRLTAFVREFLA 82
Qy 60 NGANCNLELDLNDWTALISAKSGHHIIVEELLSKGSASLEHRDGMGTALMWACYKGRDVT 119
Db 83 RGADVQAEUDDNWTALLCASRNHGLDVQVLLDHGAEEVHRDGMGTSLMWAAYRGTEL 142
Qy 120 VELLLSHGANGPSVTGLQYSYVPIIWAAGRGHADI VHL L L QNGAKVNCSDKYCTTPLVWAA 179
Db 143 VRLLLDKGADGNAHG-NYHLGALLWAAGRYKDI VELLVQRGAKVNVGKYGTITLVWAC 201
Qy 180 RKGHLECVKHLAMGADVDOEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGN 239
Db 202 RRGNEIVDTLLKAGANVDTAGYSWTPLLVAAGGHTDCVSSILEKKNVVALDKGMT 261
Qy 240 ALMIASKEGHIIVQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQYADIDIR 299
Db 262 ALCIASREGQDIAASLIAGAYINI QDRGADTPLIHAVKAGHRTVVEALLKGHADVLIQ 321
Qy 300 GQDNKTALYWEKGNATVVRDILQCNPDTEICTKDGTEPLIKATKORNI EVVELLLDKG 359
Db 322 GKDRKTAITYAVEKGHTPIVKLLLATNPDLSESATKDGDTPLLRVNRNRLIEI VHL L L DRK 381
Qy 360 AKVSAVDKDGDTPLHVAIRGRSRLAEILLRNPDKGRLLYRPNKAGETPNYDCSHQSKI 419
Db 382 AKVTASDKRGDTCLHTIAMRARSKTIIVEALLRNPKHSQLLYRANKAGETPNYDCSHQKTI 441
Qy 420 LTIQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPI CVGLIYAQWGSKGKSFLLKK 479
Db 442 LGOVFGARRLNTNDESGMLGHLYELISSALADVSEPTLTTPITVGLIYAKWGSKGKSFLLNK 501
Qy 480 LEDEMKTFAGQQTPELPQFSWLI VFL TLL L CGGLGLVFAFPV-DTNLAIAISLFLALY 538

Db 502 LRDENNFARQWABPIRTSGLLFTVCLHVALLIGTIVGLTSWAVGVSAAGFLLAY 561
Qy 539 IFFIYVYEGGREGESMNAWALSTRARHIGYLELLEFLKMFVNPPELPEOTIKALPVRF 598
Db 562 LLLAAVRYCNYQ--MDMQWASVQHGLEKRWTRLELILQVAFCHPPG-PQSDSQAKPVRF 618
Qy 599 LFTDYNRLSSVGGETSLEMIATISDACERBFGFLATLRFVFRTE--ESQGKKKKWTKC 656
Db 619 HFAEANGASPT-GDCAVAMLAALIDATESHYGWLATLYAFRPKCLKVDVGVWRWRMC 677
Qy 657 CLPSFVLEFVIGCIAGITLAF-----RVDPKHLTVNAILISIASVVGFLAVLNCR 711
Db 678 CIPVILPELALVTVTGISTVAYFTFADEKEKEHILV--ALYVIAVMGTLICHTLHV 735
Qy 712 WQVLDLSLNQRKLSAASKLHKLSEGMVKLVKCEVELMARMAKTIDSFOTQOTRLV 771
Db 736 LAKVFSVSLFTSHIRVLEAV-----RSSSAPLTMGAEVAWMTDMVKCLDAFTNQSELV 791
Qy 772 VIIDGLDACEQDKVQLMDTVRVLF--KGPIAIFASDPHIIKAINQNLSNVLRSNI 829
Db 792 GVIDALDSCDTERILTLINAVQTLTSSPNRPVLLISVDPHVIAKAAEANSRRLPTEGGI 851
Qy 830 NGHDMYRNIVHLPVFLNSRGLSNARKPLVTS-----ATNGDITCSD--TTGTQEDTDRV 882
Db 852 GGHDFLRNLHLPVYLQNSGLRKVQRAQMTALLPKRSGGQYQTDGPTLG-HSVSARRL 910
Qy 883 SQNS--LGEMTKL-----GSKTALNR-----RDYRRQWQRTITRQMSFDLTKL 925
Db 911 SNASEIISSEKLRGARGGGKRLRSESVASSTGSLNHLRQNPQTV-----LDLSRI 965
Qy 926 LVTEWFSDISPQWRRLLNIVTVGRLLRANQITFNWDLASWINLQEPYRTSMIL 985
Db 966 VLTDDYFSDVNPWRMLNVIYITVRLKAKAFIEFSWYRLSSINLQEPWPLRAIMVL 1025
Qy 986 YLEE--TECLPDQMTLKWYERIKNPIYTKDVEPLLEIDGIRNFVFLSSRTPVLAR 1043
Db 1026 HHDQFMSDNADSVLSQSVYKJRPKLAYLREAAPLLELDRDKLDAFLQHKSDLLVA 1085
Qy 1044 DVKTFLPCTVNLDPKLRRIIDVRAAREQINIGLAYPPLPHEGP-----PRPPSGYSOP 1099
Db 1086 DLRIFLPTINLDPYLRKVKEDQOTIE--DEGSLVIOARPSVNTWROFPAPTYVPSP 1143
Qy 1100 ASVCSSASFNGPPFGVVPQPHSYSGLSGQPHFYNRAAVPATGSSLL---LSSMTV 1156
Db 1144 QAYPPYQFQNEYAPN---ELSRNLSTSEPTVPLINSPS-DSFGDDILQTKLTDLTV 1198
Qy 1157 DVCEKLRQIEGLQNMWPQCTTIKKNANINGRVLSCNIDELKEHMANFGDWHLFRSM 1216
Db 1199 EGVISLLDRIBDM-KPALPKLAPVLRENAINGRVLKHCMPDLKSVLGLSFGHWELFRLL 1257
Qy 1217 VLEMRVSE-----SQVVPEDPRFLNENSSAPV-----PHGESARRS---SHTELP 1258
Db 1258 ITTLRECLRPKQORQOQOQGALEAPSNVPMIKDVTDALMQPPRESLRKNSVSHMEKQ 1317
Qy 1259 LT---ELSSQPTPTLNFSEELNTLGLDEGAPRHSNLSWQOTRTPTLSL---N 1308
Db 1318 VTLEQMICGTLQTLNEAYE-DVASSERPSTGEMLAAVAQLQOLAPRESSEFGSPDD 1376
Qy 1309 SQDSSEIETKTDKQ---ABY-RDAYREYTAQMSQL-----EGGTGSSITI---SGRSS 1355
Db 1377 QKQYGVKISNNNNNNYVLAENYNSVSHSLQSLTLVAPVGVGGGGSHLGLNGNDL 1436
Qy 1356 PHSTYVYIGQSSS-----GGSIHSTLEQ 1377
Db 1437 SDSTLDLMDHVSUVGGGGGGGTHRASRQ 1464

RESULT 15

US-10-149-819-9
; Sequence 9, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 2215706CD1
US-10-149-819-9

Query Match 25.5%; Score 2260; DB 14; Length 513;
Best Local Similarity 86.6%; Pred. No. 2.1e-147;
Matches 445; Conservative 21; Mismatches 46; Indels 2; Gaps 2;

Qy 1203 MAMNFGDHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTEL 1262
Db 1 MAMNFGDHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTEL 60
Qy 1263 SSQPTPTLNFSEELNTLGLDEGAPRHSNLSWQOTRTPTLSLNSQDSSEIETSKLTDK 1322
Db 61 SSQPTPTLNFSEELNTLGLDEGAPRHSNLSWQOTRTPTLSLNSQDSSEIETSKLTDK 120
Qy 1323 VQAEYRDAYREYTAQMSQLEGGTSGTISGRSSPHSTYVYIGQSSSGGSIHSTLQERGKE 1382
Db 121 VQAEYRDAYREYTAQMSQLEGGTSGTISGRSSPHSTYVYIGQSSSGGSIHSTLQERGKD 180
Qy 1383 GELQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDPTTEDEKSDQSGSKLLPCKKSSE 1442
Db 181 SEPKDDGRKSFMLKRGDVIDYSSGVSTNDASPLDPTTEDEKSDQSGSKLLPCKKSSE 240
Qy 1443 RPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHCSQMIKRLKAKQRECAPSQE 1502
Db 241 RSSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHCSQMIKRLKAKQRECAPSQE 299
Qy 1503 HSAEPIRTFTKAEYLSDDLDDKSDSGVRSNENSSPHSLHNEAADDSDOLEKANLTEL 1562
Db 300 HSAEPIRTFTKAEYLSDDLDDKSDSGVRSNENSSPHSLHNEAADDSDOLEKANLTEL 359
Qy 1563 EDEGHSKGRGMPHSLSLGLQDPIIARMSICEDKKSPECSLIASSPESWPAQKAYNLN 1622
Db 360 EDDSHSKGRGMPHSLSLGLQDPIIARMSICEDKKSPECSLIASSPESWPAQKAYNLN 419
Qy 1623 RTPSTVTLLNNTAPTNRANQNFDBIEGIRETSQVILRPGSPNPNTAVQNELKSMHAKRS 1682
Db 420 RTPSTVTLLNNTAPTNRANQNFDBIEGIRETSQVILRPGSPNPNTAVQNELKSMHAKRS 479
Qy 1683 QRSSTYRLSKDASLH--AASSESTGFCGERESIL 1715
Db 480 QRSSTYRLSKDPPPELHAAASSESTGFCGERESIL 513

Search completed: February 5, 2005, 19:54:49
Job time : 156 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:16:39 ; Search time 31 Seconds
(without alignments)
5322.960 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSESTGGBERSIL 1715
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5073.5	57.3	1031	2 T43458	hypothetical prote
2	1848.5	20.9	1398	2 T21884	hypothetical prote
3	1024	11.6	246	2 T46392	hypothetical prote
4	556	6.3	1881	1 SUHUK	ankyrin 1, erythro
5	555	6.3	426	2 AE2149	hypothetical prote
6	553	6.2	1856	2 B35049	ankyrin 1, erythro
7	553	6.2	1880	2 A35049	ankyrin 1, erythro
8	552	6.2	1961	2 T42716	ankyrin 3, splice
9	551.5	6.2	1940	2 T42715	ankyrin 3, splice
10	551	6.2	1943	2 T42713	ankyrin 3, splice
11	550	6.2	3924	2 S37431	ankyrin 2, neurona
12	527.5	6.0	1862	2 I49502	ankyrin - mouse
13	526	5.9	4377	2 A55575	ankyrin 3, long sp
14	516.5	5.8	1848	2 S37771	ankyrin, erythrocy
15	510.5	5.8	1765	2 T42714	ankyrin 3, splice
16	479	5.4	2039	2 T15347	ankyrin-related un
17	464.5	5.2	1549	2 T13940	ankyrin - fruit fl
18	429	4.8	2584	2 T24158	hypothetical prote
19	429	4.8	2605	2 T24157	hypothetical prote
20	414.5	4.7	1423	1 I37275	death-associated p
21	409.5	4.6	791	2 T42691	hypothetical prote
22	397	4.5	1062	2 T30255	inversin - mouse
23	384.5	4.3	557	2 T46507	hypothetical prote
24	382	4.3	1062	2 T14151	Inv protein - mous
25	378.5	4.3	323	2 B47169	ankyrin-like repea
26	361	4.1	397	2 T46445	hypothetical prote
27	356	4.0	679	2 B45771	2-5A-dependent RNA
28	347.5	3.9	741	2 A45771	2-5A-dependent RNA
29	338.5	3.8	991	2 T25412	hypothetical prote

30	335	3.8	1401	2 S11527	alpha-latrotoxin p
31	325	3.7	1188	2 T19552	hypothetical prote
32	322.5	3.6	441	2 S74626	erythroid ankyrin
33	321.5	3.6	934	1 H71274	probable ankyrin -
34	302	3.4	1058	2 D82654	ankyrin-like prote
35	301	3.4	2004	2 D88948	ankyrin-like prote
36	300.5	3.4	1411	2 S30355	alpha-latrotoxin
37	296.5	3.3	627	2 E86190	hypothetical prote
38	294	3.3	1184	2 T00253	gene Ankhzn protoi
39	292	3.3	606	2 AC2508	hypothetical prote
40	288	3.3	1083	2 S48460	probable membrane
41	287	3.2	517	2 T48283	ankyrin-like prote
42	286.5	3.2	662	2 E84725	ankyrin-like prote
43	285	3.2	1001	2 S30385	G9a protein - huma
44	279.5	3.2	368	2 T18184	ankyrin repeat pro
45	279	3.2	543	2 C86212	hypothetical prote

ALIGNMENTS

RESULT 1

T43458
hypothetical protein DKFZp434F0621.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43458
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z22517
A:Accession: T43458
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1031 <AAA>
A:Cross-references: UNIPROT:Q9UF42; EMBL:AL133620
A:Experimental source: adult testis; clone DKFZp434F0621
C:Genetics:
A:Note: DKFZp434F0621.1

Query Match 57.3%; Score 5073.5; DB 2; Length 1031;
Best Local Similarity 95.2%; Pred. No. 8.1e-291;
Matches 974; Conservative 23; Mismatches 13; Indels 13; Gaps 1;

Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN	60
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKN	60
Qy	61	GANCNLEDLDNWTALISASKEGHIHVEELKSGASLEHRDMGGWTALMWACYKGRD	120
Db	61	GANCNLEDLDNWTALISASKEGHVHVEELLKCGVNLHRDMGGWTALMWACYKGRD	120
Qy	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLV	180
Db	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLV	180
Qy	181	KGHECVKHLAMGADVDGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGN	240
Db	181	KGHECVKHLAMGADVDGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGN	240
Qy	241	LMTASKEGHIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIR	300
Db	241	LMTASKEGHIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIR	300
Qy	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATMRNIEVVELLDKGA	360
Db	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATMRNIEVVELLDKGA	360
Qy	361	KVSADKKGTPLHVAIRGSRRLAEILLRNPKDGRLLYRNKAGETPYNDCSHQSKIL	420
Db	361	KVSADKKGTPLHVAIRGSRRLAEILLRNPKDGRLLYRNKAGETPYNDCSHQSKIL	420
Qy	421	TQIFGARHLSPTDGDMLGYDYLSALADILSPTMQPPTCVGLYAQWGSCKSFLLKKL	480
Db	421	TQIFGARHLSPTDGDMLGYDYLSALADILSPTMQPPTCVGLYAQWGSCKSFLLKKL	480

Db 154 QGHENVVAHLINYGKGVLPALHIAARNDDTAAVLLQNDPNPDVLSKGTGTPHIA 213
QY 84 -----THIVEELKSGASLEHRDMGWMTALMAWACYGRDVTVELLSHGANPVSVTGLQYS 138
Db 214 AHYENLVAQLLNRGASVNTFQNGITPLHIAARRGNVIMVRLLDGRGAQIE-TTKDE 272
QY 139 VYPIIWAAGRHADIHLLQNGAKNCSDKYGTTPVWAARGHLECVKHLAMGADV 198
Db 273 LTPHCAARNGHVRISEILLDHGAPQAKTNGLSPIHMAAGDHLDCKVLLQYDAEID 332
QY 199 -----QEGA-----NSMTALIVAVKGYTOSVKEILK 225
Db 333 DITLDHLTPHVAACHGHRVAKVLLDKGAKPSRALNGFTPLHIAKKNHVRVMEILLK 392
QY 226 RNPVNLTDKGNMTALMIASKEGHEIIVQDLIDAGTVNIPDRSGDTVLIGAVRGHVEI 285
Db 393 TGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSKVETPLHMAARAGHTEV 452
QY 286 VRALLQKYADIDIRGDNKTALYAVEKGNATVVDILOCNPDTEICTKDGETPLIKATK 345
Db 453 AKYLLQNKAKNAKADQDPLHCAARIGHTNNVKKLLENANPNLATTAGHTPLHIAAR 512
QY 346 MRNIEVVELLDGAKYSADVKKGDTPHVAIRGRSRLAEILLRNPKDGRLLYRPNKAG 405
Db 513 EGHVEITVALLKEASQACMTKKGFTPLHVAKYGVKRVVAELLE-RDA----HENAAG 566
QY 406 E-----TPYNIDCSHQ-----KSIL-----TQIPGAR----- 427
Db 567 KNGLTPLHVAHVHNNLDIVKLLPRGSGSPHSPAWNGYTPHIAAKQNVQEVARSLLQYGG 626
QY 428 -----HLSPTEDGMLGYDLYSSALADILSEPTWOP-----PIC-- 462
Db 627 SANAESVQGVTPHLHAAQGAHAEKVALLSKQNGNLGNKSGLTPLHVAQEGHVPADV 696
QY 463 -----VGLYAQWGSKG--SFLKKLED-EMKTFAGQOOTEPLFQFSWL 501
Db 687 LIKHGVMVDATTMGYTPHVAHYGNIKLVKELLQHQADVNAKTKLGY--SPHQAQAQ 744
QY 502 --IVFLTLLCGGLGVFPADV-----TNLAIAISLFLALIIYFIVYIGGRERGE 553
Db 745 GHTDITVLLKNG-----ASPNEVSSDGTTPLAIAKRLGYISVTDVLKV----- 789
QY 554 SNWMAWLSRLARHIGYLELLFKLMFVNPPELPEQTKALPVRELFDTNRLS-SVGG 612
Db 790 TDTSFVLVSDKIR-MSFPTVDEILDVSEDEGELIS-----FKAERRSRVDDEE 840
QY 613 TSAEMIATLSDACEREFGLATRLPR-----VFRTEE-SQKKKWKTKCCLPSFV 662
Db 841 KELLDFVPKLDQVVE-----SPAIPRIPCAMPETVIRSEEQEASKEYDEDSLIPS-- 892
QY 663 IFLFIVGCIITAGITLAIFRVDPKHLTVNALISIASVUGLAVFLNCRTWQVLDLSLNS 722
Db 893 -----SPATETSDNI-SPVASPVHTGFLVSP-----MVDARGGS 925
QY 723 QRKRLHSAASKHLKLSGKPMVKCEVELMARMAKTIIDFTQNTQRLVLIIDGLDACEQ 782
Db 926 MRGSRHN-----GURVIP-----PRTCAAPTITCLV----- 954
QY 783 DKVLQMLDVTVRVLFSGPPIAFASDPHIIKAINQNLNSVLNRDSNINGHYMRN-IVHL 841
Db 955 --KPQKLSLTPPL-----ABEEGLASRIIALTGPTGAQFLSPVIVEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVQNSLGENTKLGSKTALN 900
Db 994 PFAHSHRG---DRELVLVLRSENGV-----WKEHRSRYGESYLDQILN-GMDELG 1041
QY 901 RRDYTRRRMORTITROMSFDLTKLVT--EDWFSDISPTQWRLLNINIVSVTGRLLRAN 957
Db 1042 SLELEKRVCRIT--TDPLFVINSRLCQD-YDIIGPEG-----GSLKSLVPLV 1091
QY 958 QITFNMDRLASWINLTQWPYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005

Db 1092 QATFPENAVTKRVKLAQ-----AQVPVDELVTKLGNQATFSPIVTVPR 1137
QY 1006 ISK-----NIPTKDVEPLEIDGDIRFEVF-----LSRTPVLVA 1042
Db 1138 RRFHPIGLRIPLPPSWTNDPRDSGEDTTLRLLCVSGVIGTDAQMEDITTKLVYA 1197
QY 1043 RDVKTEPLCTVNLDPKRLRIIADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSPASV 1102
Db 1198 NECANF--TTNV--SARFWSDCPRTAAVNFATILLYKEL-----TAVPYMAKF 1242
QY 1103 CSASFNPGPPGWSVPQPHSSYYSGLSGQHPFYFNRAAVPATGSLSSMTVDVCEK 1162
Db 1243 VIFAKMNDPREG-----MAMNFGD-WHLFRSMVLEMRSVESQVVPEDPR----- 1269
QY 1163 LRQ-----LEG-----LDQNMVPOCTTIKAN-----INGRVL 1191
Db 1270 LEQHENFVARSRDIEVLEGMSLFAELSGNLVP-----VKAAQORSFHQSFRENRLA 1324
QY 1192 SQCNIDELKE-----MAMNFGD-WHLFRSMVLEMRSVESQVVPEDPR----- 1233
Db 1325 MPVKVDRDSREPGGSLFLRKAMKYEDTOHILCHLNIWPPCAKSGAEDRRRTPTPLAL 1384
QY 1234 ---FLNENSAPVPHGESARR-----SSHTEPLTESSQTPYTLNPFSEELNTGLD-- 1283
Db 1385 RYSLSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARE---LQFSVEDINRIRVENP 1440
QY 1284 ---EGAPHSNLSWOSQTRTPSLSSINSQSSSIEISKLTQVQA-----EYRD 1329
Db 1441 NSLLEQSVALLNL-WYIREGONANMENLYALQSIDRGEIVNNLEGSGRQSNLKPDRRH 1499
QY 1330 AVREYIAQMSQLEGGTISGRSSPHSTYVYIGQSSGSIHSTLEQREGKELQED 1389
Db 1500 TUDYDLSLSPQNG--YSSLQDELSP-----ASLGCALSSPLRADQ----- 1539
QY 1390 GRKSFMLKRGDVIDYSSSGSVNNEASPLDPI-----TEED---EKSQOS--GSKLLPCKKS 1440
Db 1540 -----YMNVAVLDAIPLAATEHDTMLEMSDMQVMSAGLTPSLVT 1579
QY 1441 SERPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVQITPHCSKWIRTKLKAQRECAP 1500
Db 1580 AEDSSL-----ECSKAEDSDATGH-----EWKLEGALESEPRGP 1613
QY 1501 QEHSABPIRTFIKAKEYLSDA---LLDKKSDSDSGVRSNESPNNHSLHNEAADDQLEKA 1557
Db 1614 ELGSLR---LVEDDVTVDSDATNGLIDLLE-QEEOGORSSEKLPFGSKRQDDATGAGQ-DSE 1667
QY 1558 NLIELEDEGHGKRGMPHSLGSLQDPIIARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1668 NEVSLVSGHQRGQARITHS-----PTVSQVTERSQDRLQDWDADGSIYSYLQDAAGSW 1721

RESULT 5

AE2149

hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AE2149

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; WUID:21595285; PMID:11759840

A;Accession: AE2149

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-426 <KUR>

A;Cross-references: UNIPROT:Q8YTG9; GB:BA000019; PIDN:BA074447.1; PID:g17131841; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all2748

Query Match 6.3%; Score 555; DB 2; Length 426;

Best Local Similarity	37.2%	Pred. No. 5.6e-25;
Matches 149; Conservative	67;	Mismatches 177; Indels 8; Gaps 6;
Qy	9	VINTVEENIPAL ¹ KALLEKCKVD ¹ ERNECG ¹ OTPL ¹ MLAAEOGN ¹ EV ¹ VELLKNGANGCN ¹ L-E 67
Db	8	LLK ¹ VAKSGD ¹ IKGL ¹ GALLA ¹ AGVGV ¹ DI ¹ CRD ¹ GT ¹ TAL ¹ FA ¹ NLGY ¹ TEIV ¹ RSLLDGGANV ¹ NLAR 67
Qy	68	DLDNWT ¹ TAL ¹ ISASKEGH ¹ IH ¹ VEEL ¹ LKSGASLEH ¹ RD ¹ MGW ¹ TAL ¹ MWAC ¹ YKGR ¹ TDV ¹ VELL ¹ SHG 127
Db	68	KRYGL ¹ TAL ¹ MLAAS ¹ ANQ ¹ VD ¹ VL ¹ LIS ¹ RGA ¹ AVNAT ¹ NEG ¹ STAL ¹ MAA ¹ LKGN ¹ VEAR ¹ VLLA ¹ AG 127
Qy	128	ANPSVT ¹ GLQ ¹ SVY ¹ PI ¹ IWA ¹ ARGH ¹ ADI ¹ VH ¹ LLQ ¹ NGAK ¹ VNC ¹ SD ¹ KY ¹ GT ¹ TP ¹ LWA ¹ ARK ¹ HUE ¹ CV 187
Db	128	ADVNIT ¹ DKDD ¹ T-AL ¹ KLAV ¹ RKG ¹ AAV ¹ QL ¹ ILP ¹ SGADAN ¹ CEDE ¹ GET ¹ TL ¹ MLA ¹ ADSG ¹ HGDV ¹ 186
Qy	188	KHLLAMGAD ¹ VD ¹ OGANS ¹ WTAL ¹ I ¹ VA ¹ VKG ¹ Y ¹ Q ¹ SVKE ¹ IL ¹ KRN ¹ PNV ¹ NLT ¹ DKD ¹ GN ¹ TAL ¹ MIASKE 247
Db	187	QVL ¹ LAAGV ¹ DVNE ¹ QD ¹ GGT ¹ ALLA ¹ AVAG ¹ NAI ¹ AKILL ¹ DRGADV ¹ NHQD ¹ GESAL ¹ HLAT ¹ VE 246
Qy	248	GHIEIV ¹ QDL ¹ LAD ¹ AG ¹ TYVNI ¹ PDR ¹ SGD ¹ TVL ¹ IG ¹ ARG ¹ GHVEI ¹ VRALL ¹ QYAD ¹ IDIRG ¹ D-NKTA 306
Db	247	GYVDV ¹ VQVLL ¹ NOGANT ¹ QIK ¹ KL ¹ GT ¹ PL ¹ VAAL ¹ QHDQ ¹ IVETLL ¹ -KYG-ANVH ¹ GNL ¹ GETP 304
Qy	307	LYNAVEKGNAT ¹ WVRD ¹ IL ¹ QCNP ¹ DT ¹ EICT ¹ KD ¹ GET ¹ PL ¹ IKAT ¹ KMRNIEV ¹ VELL ¹ LK ¹ GA ¹ KYSAVD 366
Db	305	LTLAASQ ¹ GHAT ¹ VRILL ¹ DYGANANI ¹ PASD ¹ CKT ¹ ALI ¹ KATERNH ¹ PCVI ¹ QILL ¹ LKAGANV ¹ YQD 364
Qy	367	KKGD ¹ TP ¹ LHVAIR ¹ GRSRRL ¹ ABL ¹ LLRN ¹ PKD ¹ GRLL ¹ YRPNKAGET 407
Db	365	SVGATALI ¹ WAASGYN ¹ KVVOILL ¹ EGGAD ¹ TNL---KNRGGYT 402

RESULT 6
B35049
ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1856 <LAM>
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: Spl1.2-8p11.2
C;Superfamily: ankyrin repeat homology
C;Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;3-1513.1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>

Db 926 MRGRHN-----GLRVIP-----PRTCAAPRITCELV----- 954
 Qy 783 DKVLQMLDTRVLFSGKPFIAFASDPHIIKAINQNLNSVLDSNNGHYNRN-IVHL 841
 Db 955 --KPQKLSPPPL-----AEEGLASRIITAGTGAQFLSPVIVEI 993
 Qy 842 PVFL-NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRVSQNSIGENTKLGSKTALN 900
 Db 994 PFAFHSGRG---DRELVLVSESGV-----WKEHRSRYGESYLDQILN-GMBEELG 1041
 Qy 901 RRDYTRRQMRITRQMSFDLTKLLVT---EDWFSDISPQTMRLNLINIVSVTGRLLRAN 957
 Db 1042 SLELEKRVCRIT--TDPLFVMSRLCQD-YDIIGPEG-----GSLAKSLVPLV 1091
 Qy 958 QITFNWDRLASWNLTEQWYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005
 Db 1092 QATFPENAVTKRVKALQ---AQPVDELVTLLGNQATFSPVITVEPR 1137
 Qy 1006 ISK-----NIPTKDVEPLEIDGDIRNEVF-----LSRTPVLVA 1042
 Db 1138 RRFPHRFIGURILPFSWTDNPRDSGEGDTTSLRLCSVTGGTDQAWEDITGTLVYA 1197
 Qy 1043 RDVKTFPCTWNLDPKLREIADVRAAREQINIGGLAYPPLPLHEGPPRPSPGVSQPASV 1102
 Db 1198 NECANF---TTNV--SARFWSLSDCPTAEAVNATLLYKEL-----TAVPYWAKF 1242
 Qy 1103 CSSASFNGPPFGVSPQSHSYVGLSGPQHPFYRNAAPVATGSSLLLSMTVDVVECK 1162
 Db 1243 VIFAKMDPREG-----ANINGRVLSCQNIDELKKEMANFGDWHLPFS--M 1216
 Qy 1163 LRQIEGLDQNMPOYCTTIKK---ANINGRVLSCQNIDELKKEMANFGDWHLPFS--M 1216
 Db 1270 LECHENFEVARSRDIEVLEGMSLFAELSG-----NLVPVKAQAQORSFHQSFRENRL 1323
 Qy 1217 VLEMRSVESQVP-----EDPRFLNENSAVPV-----HGESARRSSHTLP 1258
 Db 1324 AMPKVYDSSREFGGSLSPLRKAMKYEDTQHILCHLNIWPPCAKSGAEDRRTPPTPLA 1383
 Qy 1259 L--TELSSQTPYTLNFSE-ELNLTGLDEGAPHNSLNWSQSTRTTP-SLSSLSNSQDSI 1314
 Db 1384 LRYSLSESTPGSLSGTEQAEKMAVISE---HLGLSWAELARELOFSVEDIN----- 1433
 Qy 1315 EISKLTQVQAEYDAREYIAQMSQLEGGTSGSTISGRSPHSTYVI--GQSSSGGSIH 1372
 Db 1434 -----RIRVENPNSLLEQSVALLNL-----WVIREGQANMENLY 1468
 Qy 1373 STLEQ-ERKGEGLKQDGRKSLFKM---RGVDIYSSGVSVTNEASPLDPITEEDR--- 1425
 Db 1469 TALQSIDRGEIVNVLGSGRQSNLKPDRRHTDRDYSLSPSQMGYSLL-----QDELLS 1523
 Qy 1426 -----KSDQSGSKL-----LPGKKSSEPSLFTQDLKLKGGGLRYQKLPSDED 1468
 Db 1524 PASIGCALSSPLRADQYWEVAILDAIPLAATEHDTMLEMSDMQVWSAGLT-PSLVTAED 1582
 Qy 1469 ESGTGRVQIIPHCSEKMTKELKAKQECASQBSAEPTRTFKAXEYLSDALLDKDS 1528
 Db 1583 SS-----LECSK---AESDATHENKLEGALSEEPGPELSELVEDDITVDSGAT 1631
 Qy 1529 S-----DSGVRNSSESSPNHSLHNEAADDQSLKANKANLIELEDEHSGRKGMPHSLSL 1580
 Db 1632 NGLDILLEQEBGQSEKLPQSKQDDATGAGQ--DSENEVSLVSGHQRGOARITHS----- 1686
 Qy 1581 QDPIIARWICSEDK-----KSPSECSLIASPPESW 1612
 Db 1687 --PTVSQVTRSQRDLQWDADGSIYSYLODAAQGSW 1721

RESULT 7

A35049
 ankyrin 1, erythrocyte splice form 2 - human
 N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
 N;Contains: ankyrin 2.2, erythrocyte
 C;Species: Homo sapiens (man)

C;Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 09-Jul-2004
 C;Accession: A35049
 R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; Ke
 Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
 A;Title: cDNA sequence for human erythrocyte ankyrin.
 A;Reference number: A35049; MUID:90175370; PMID:1689849
 A;Accession: A35049
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1880 <LAM>
 A;Cross-references: UNIPROT:P16157; GB:M28880
 C;Genetics:
 A;Gene: GDB:ANK1; ANK
 A;Cross-references: GDB:118737; OMIM:182900
 A;Map position: 8p11.2-8p11.2
 C;Superfamily: ankyrin; ankyrin repeat homology
 C;Keywords: alternative splicing; cytoskeleton
 F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
 F;2-1513-1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
 F;44-76/Domain: ankyrin repeat homology <AN01>
 F;77-109/Domain: ankyrin repeat homology <AN02>
 F;110-142/Domain: ankyrin repeat homology <AN03>
 F;143-171/Domain: ankyrin repeat homology <AN04>
 F;172-204/Domain: ankyrin repeat homology <AN05>
 F;205-237/Domain: ankyrin repeat homology <AN06>
 F;238-270/Domain: ankyrin repeat homology <AN07>
 F;271-303/Domain: ankyrin repeat homology <AN08>
 F;304-336/Domain: ankyrin repeat homology <AN09>
 F;337-369/Domain: ankyrin repeat homology <AN10>
 F;370-402/Domain: ankyrin repeat homology <AN11>
 F;403-435/Domain: ankyrin repeat homology <AN12>
 F;436-468/Domain: ankyrin repeat homology <AN13>
 F;469-501/Domain: ankyrin repeat homology <AN14>
 F;502-534/Domain: ankyrin repeat homology <AN15>
 F;535-567/Domain: ankyrin repeat homology <AN16>
 F;568-600/Domain: ankyrin repeat homology <AN17>
 F;601-633/Domain: ankyrin repeat homology <AN18>
 F;634-666/Domain: ankyrin repeat homology <AN19>
 F;667-699/Domain: ankyrin repeat homology <AN20>
 F;700-732/Domain: ankyrin repeat homology <AN21>
 F;733-765/Domain: ankyrin repeat homology <AN22>
 F;766-798/Domain: ankyrin repeat homology <AN23>
 Query Match 6.2%; Score 553; DB 2; Length 1880;
 Best Local Similarity 20.5%; Pred. No. 9e-24;
 Matches 388; Conservative 287; Mismatches 648; Indels 574; Gaps 73;
 Qy 21 LKALLEKCDVDRNECGOTPLMAAEQGNVEIVKELLKNGKNCNLEDLNDWTALISASK 80
 Db 94 VRELNVYGVANVNAQSQKGFPLYMAAQENHLEVVFLENGANQNVATEDGFTPLVALQ 153
 Qy 81 EGH----- 83
 Db 154 QCHENVVAHLINYGTKGVRLPALHIAARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIA 213
 Qy 84 -----THIVEELKSGASLEHEDWGTALMAWACYKRTDVELLSHGNPSVTGLQYS 138
 Db 214 AHYENLVNVAOLLNRSSVNFPTQNGITPLHISRRGNVIMVRLLLDRGAQIE-TTKDE 272
 Qy 139 VYPIIWAAGRGHADIVHLLLQNGAKVNCSDKYGTTPPLWMAARKHLECVKHLAMGADV 198
 Db 273 LTPHCAARNGHVRISEILLDHGCAIQAKTKNGLSPIHMAAQGDHLDCCVLLQYDAID 332
 Qy 199 -----QEGA-----NSMTALIVAKGTYQSVKEILK 225
 Db 333 DITLDHLTPLHVAACHCHHRVAKVLLDKGAKPNRSALNGFTPLHIAACKNHRVMEILLK 392
 Qy 226 RNPVNLTKDGNLTALMIASKEGHEIIVODLLDAGTVVNPDRSGDVTVLICAVRGHVEI 285
 Db 393 TGAISDAVTESGITPLHVASFMGHLPIVKNLLOGRASPNSVNVKVEFPLHMAARAGTEV 452
 Qy 286 VRALLQKYADIDIRGQDNKTALYAWVEKGNATVVRDILQCNPPTEICTKDGETPLIKATK 345

Db 453 AKYLONKAKNAKADQDTPLHCAARIQHTNMVKKLLNNANPNLATTAGHTPLHIAAR 512
QY 346 MENIEVELLDKGAUVKADKKGTOTPLHVAIRGRSRLAEALLRNPKDGRLLYRNKAG 405
Db 513 EGHVETVLALLEKASQACWKTKGFTPLHVAKYKVRVAELLLE-RDA-----HPNAG 566
QY 406 E---TPYNIDCSHQ-----KSTL-----TQIFGAR----- 427
Db 567 KNGLTPLHVAVHNHNLIDVKKLLPRGSPHSPANNGYTPHLIAAKONQVEVARSLQYGG 626
QY 428 -----HLSPTETDGMGLDYLYSSALADILSEPTMQP-----PIC-- 462
Db 627 SANAESVOGVTPLHAAQGAHAEVALLSLKQANGNLGNKSGLTPLHLVAQEGHVPVADV 686
QY 463 -----VGLYAOWGSGK--SFLILKLEL-EMKTFAGQOTEPLFQFSWL 501
Db 687 LIKHGVMDATRMGYTPPLHVAHYGNIKLVKFLLOHQADVNAKLYG--SPLHQAAQQ 744
QY 502 --IVPLTLLCGGLGLVFAFPVD-----TNLAIAISLFLALYIPFIVYFGRREGE 553
Db 745 GHTDVTVLLKNG-----ASPNEVSSDGTTPLAIAKRLGYISVTDVLKV-- 789
QY 554 SNNAWALSTRLARHIGYLELLFKLMFVNPPELPOTKALPVRFDTYRLS-SVCGE 612
Db 790 TDTSFVLVSDKHR--MSFPETVDEILDVSEDEGELIS-----PKAERDRSDVDEE 840
QY 613 TSLAMIATLSDACERBEGLATLFR-----VETEE-SQKKKKWKTCCLPSPV 662
Db 841 KELLDVFKLQOVB-----SPAIPRCPAMPETVIRSEEQASKEYDEDSLIPS-- 892
QY 663 IFLFVGCIIAGITLAIIFRVDPKHLTVNAILISIASVVLGFLVNLNCRMTQVLDLSLNS 722
Db 893 -----SPATETSDNI-SPVASPVHTGELVSF-----MVDARGGS 925
QY 723 QRKRLHSAASKLHLKSGEFMKVLCCEVELMARWAKTIDSTQNTQRLVLIIDGLDACEQ 782
Db 926 MRGSRHN-----GLRWIP-----PRTCAAPTRITCRV----- 954
QY 783 DKVLQMLDTRVLVFSKGFPIAFASDPHIIKAINONLNSVLNRDSNINGHDYMN-IVHL 841
Db 955 --KQKSLTPPL-----ABEGLASRIIALGPTGAOFLSPVIVEI 993
QY 842 PVFL-NSRGLSNARKFLVTSATNGDITGSDTTGTQEDTDRRVSONSLGEMTKLASKTALN 900
Db 994 PHFASHGRG---DRELVLVRENGSV-----WKEHRSYGESYLDQILN-CMDEELG 1041
QY 901 RRDYRRQMQRTTROMSFOLTKLLVT---EDWFSIDISPTMRRLNIVSVTGRLLRAN 957
Db 1042 SLELEKKRVCRITIT--TDFFLYFVIMSRQCQD-YDIIGPRG-----GSLKSKLVPLV 1091
QY 958 QITFNWDELASWINLTQWPYRTSWLILYLEETEGLPDQMTLK-----TWYER 1005
Db 1092 QATFPENAVTKRVKALQ-----AQVFPDELVTKLGNQATFSPITVTEPR 1137
QY 1006 ISK-----NIPTTKQVPELLEIDGDIRNFEVF-----LSSRTPVLVA 1042
Db 1138 RRKHFRPGLRILPFPSTWPNRPSGEGDTTSLRLCSVIGGTQAOQWEDITGTTKLVA 1197
QY 1043 RDVKTFLPCTVNLDPKLEIADVRAAREQINIGLAYPLPLHEGPPRPPSGYSOPASV 1102
Db 1198 NECANF---TTNV--SARFMLSDCPRTAEAVNFTLLYKEL-----TAVPYMAKF 1242
QY 1103 CSSASFNGPPPGVGVSPQPHSSYSYSGLSGPHQPFYRNAVPATGSSLLLSMTVDVCEK 1162
Db 1243 VIFAMNDPREG-----RLRCYCMTDKDKYKT 1269
QY 1163 LRQTEGLDQNMPOYCTTIKK-----ANINGRVLSQCNIDELKKEMANFGDWHLFRS--M 1216
Db 1270 LEQHENFVEVARSDIEVLEGMSLFAELSG-----NLVPVKAAQORSFHFQSFRENRL 1323
QY 1217 VLEMRSVESQVPP-----EDPRFLNENSSAPVP-----HGESARRSSHTPLP 1258
Db 1324 AMPVKVRDSSREPFGGSLFLRKAMKYEDTQHILCHLNITWPPCAKSGAEDRRRTTPLA 1383

QY 1259 L--TELSSQTPVTLNFSPE-ELNTLGLDEGAPRHSNLSWQSOTRTP--SLSSLSQDSSI 1314
Db 1384 LRYSLSESTPGSLSGTEQAEWKMAVISE-----HLGLSWAELARELQFSEVDIN----- 1433
QY 1315 EISKLTQVQAEYDAYREYIAQMSQLGEGGTGSGSTISGRSSPHSTYYI--QOSSSGGSIH 1372
Db 1434 -----RIRVENPNSLLEQSVALLNL-----WVIREQONANWENLY 1468
QY 1373 STLBO-ERKKEGELKQEDGRKSPFLMK---RODVYDYSVSGVSTNEASPLDPITEDE-- 1425
Db 1469 TALQSIDRGEIVNMLEGSGRSLNPKDRRHTDRYDLSPSQMGYSYL-----QDELLS 1523
QY 1426 -----KSDQSGSL-----LPGKKSRRPSIFQTDLLKGGGLAYQKLPSPED 1468
Db 1524 PASLGCALSSPLRADQTNVEVAIILDAIPLAATEHDTMLEMSDMQVWSAGLT-PSLVTAED 1582
QY 1469 ESGTGRVQITPHCSKMIRTKRLKAKQREKASPOHSAEPIRTFKAKLEYLSDALLDKKDS 1528
Db 1583 SS-----LECSK---AEDSDATGHEWLEGALSEEPGPELGSLELVEDDTVDSAT 1631
QY 1529 S-----DSGVRSNESSPHLSLNEAADSQLEKANLIELEDEHSGKMGPHSLSGL 1580
Db 1632 NGLIDLLEQEBQGRSEKLPGRKQDDATGAGQ--DSENEVSLVSGHQRGQARITHS----- 1686
QY 1581 QDPILARMSICSEDK-----KSPSECSLIASSPEESW 1612
Db 1687 --PTVSQTVRSQRLQDMADGSIYSVYLODAAQGSW 1721

RESULT 8

T42716

ankyrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C:Accession: T42716

J. Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633; PMID:7615634

A:Accession: T42716

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1961 <P>

A:Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:G710548; PID:G710552; PIDN:AA01607

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 6.2%; Score 552; DB 2; Length 1961;

Best Local Similarity 20.2%; Pred. No. 1.1e-23;

Matches 427; Conservative 289; Mismatches 674; Indels 724; Gaps 85;

QY 14 EENIPALKALKKCKVDVERNECGQTPLMAAEEQGNVEIVKELLKNGANCNLEDLQNW 73

Db 66 KEGHVEVVELQREANVDAATKKNTALHSLAGQAEVVKLVITNGANVNAQSQNGFT 125

QY 74 ALISASKEGHIHVEELKSGASLEHRDMCGWTALMWACYKGRDVTVELL----- 124

Db 126 PLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQQHQDVVSLLENDTKGVR 185

QY 125 -----SHGANPSVTGI----- 135

Db 186 LPAHIAARKDDTKAAALLLQNDTNADVSKSGFTPLHIAAHYCNINATILLNRAA 245

QY 136 ---QYSYVPIIWAAGRHADIVHLLQNGAKVNCDSKYGTTPLVWAARKG----- 182

Db 246 FTARNDITPLHVASKRGNANWVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVVELLD 305

QY 183 -----HLECVKHLAVGADYDOEGANSMTALIVAKGVYQTS 219
Db 306 RSAPILSKTKNGLSPLHMATQGDHLNVCQLLOHNVPVDDVTNDYLTALHVAACHGYKV 365
QY 220 VKEILKRNPNVNTDKDGNALTAMASKE----- 247
Db 366 AKVLLDKASPNKALNGFTPLHLACKNRIWVWELLKHGASIQAVTESGLTPIHVAAF 425
QY 248 -GHIEIVQDILLDAGTVYVNPDRSGDTVLICAVRGHVEIVRALLQKVADIDIRQONKTA 306
Db 426 MGHVNIQSOLMHHGASNTTNVGETALHMAARSGQAEVRYLVQDQAQVEAKAKDQTP 485
QY 307 LYNAVEKGNATMRDILQCPDTEICTKQGETPLIKATKRNIEVVELLDKAKYSADV 366
Db 486 LHSARLGKADIQOOLQOGASPNAAATTSYTPPLHLAAREGHEDVAFLDHCASUSITT 545
QY 367 KKGDTPLHVAIRGSRRLAELLAR-----NPKDGRLL----- 398
Db 546 KKGFTPLHVAKYCKLEVASLLQKSASPDAGKSGLTPLHVAHYDNQKVALLLDQGA 605
QY 399 --YRPNKAGETPNIDC-SHOKSILTOI--FGAR-----HLSPTETDGMGLGY 441
Db 606 SPHAAKNGYTPHLIAAKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDMSVL 665
QY 442 DLYSSALADILSEPTWOP-----PICVGLYAOW 469
Db 666 LLSRANVNLSNKSGLTPLHLAQEDRVNVAEVLVNOGAHVDAQTKMGYTPPLHVGHCH--Y 723
QY 470 GSGK--SFLIK--KLEDEMKTFAGQOQTEPLFQSMILVFLTL-----LLCGGLGLV 516
Db 724 GNKIVNFFLOHSAKNVAKTNGVYALHQAQQGHTHIINVLLQNNASPNELTVNG--- 779
QY 517 FAPPVDTNLAIAISLSFLALIIYFFIVI----- 544
Db 780 -----NTALAIARRLGIVVDTLKVTEEIMTTTTITEKHKMNVPEWNEVLDMSDEV 834
QY 545 -----YFGRREGESNNAWALSTRLARHIGYLELLFKL-----MFVN 582
Db 835 RKASAPKLSGGEYISDGEQKCTWPKIPKVQ-----EVLVSEDAITGDTKYL 886
QY 583 PPPELPEQTKALPVRFLEFDYNNRLSSVGGTSLAEMIATILSDACERFEGFLATRLFRVFR 642
Db 887 PQDLKELGDDSLPAE-----GYVG--FSLGARSASLR-----SFSSDRSYTLNR 928
QY 643 TBSQKKKKKKKCCCLPSFVIFLEIVGCIITAGITLLAI PRVDP-KHUT-----VNAI 693
Db 929 SSVAR-----DSMMIEELLVPSKEQHLTFTRFDSLSLRHYSWAADTLDNVNLV 977
QY 694 LISTASVVLGAFVLCNRTWQVLDLSLNSQ-----RK-----RLHSAASKLHKL- 737
Db 978 SSPVHSGFLVSFWVDARG-----GSMGRSRHGNRIIIPPRKCTAPRITCRLVVRHKL 1032
QY 738 -----KSEG-----FMKVLKCEVELMARMAKTIDSFOTN 766
Db 1033 NPPWVSEGLASRLVEMGAPAGFLGPVIVEIPHFGMRGKERELIVLRSENGETWKEH 1092
QY 767 Q-----TRLVVIDGLDACEQDKVQLMLDTRV--LFSKGPFTIAFASDPIHIIKAINQN 819
Db 1093 QFDSKNEDLAEELNGMD--BELDSPEELGTKRICRIITK-DFOQYFA-----VVSRIKOE 1144
QY 820 LNSVLRSNNGHDYMENIVHLPVFLNSRGLSNARKFLV-TSATNGDITCSDTTGTG--- 875
Db 1145 SNQIGPSEGGI-----LSSVTFLVQASFPPEGATKRIRVGLQAOP 1184
QY 876 --EDTDRRVSONSLGEMTKLGSKTALNRDITY--RRRQMORTITROMSFOLTKL-----VT 928
Db 1185 VPEETVKKI-----LGNKATFSPIVTEPRRRKFKHPiF--MTIPVPPSGEGVS 1232
QY 929 EDWFSDSIPQMRRLINIVSVTG-----RLLRANQIIFPNWRDLASWINLIEQWYPT 980
Db 1233 NGYKGDATPNL--RL--CSITGTSPAQWEDITGTTPLTFIKDCVGFNTNVSARF--- 1284
QY 981 SWL-----ILYLEETEGLPDQWTLKTM-YERISKNIPTTKDVEPLLEIDGDIRNF----- 1029

Db 1285 -WLADCHQVL---ETVGLASQLYRELI CVPYNAKFVFAKTNDP-----VSSSLRCFCWTD 1337
QY 1030 -----EVLSSRTPLVAVARDVKTELPCTVNLDPKLR-----EIIADVAAARE-- 1071
Db 1338 RVDKLTLEQOENFEAVARSKDIEVLEKPI--YVDCYGNLAPLTGKGQQLVFNFSYSPKENR 1395
QY 1072 ---QINIGGLAYPPLP---LHEGPPRPFGSGYSPASVCSSASFNPGFPGVVPQPQPHSS 1124
Db 1396 LPSIKIRDTSOEPCGRLSFLKE--PKTTKGLPQTA--VCN----- 1432
QY 1125 YVSGLSGPOHPFNRAAVPATGSSILL-----SMTVDVVCEK--LRQIEGLDQNM 1173
Db 1433 --LNIITLPAHKAEXADRQSFASLALRKRYSLTEPMSQSPQSCERTDIRMAIVAD-HL 1489
QY 1174 MPOYCTTIKANINGRVLSQCNIDELKEMAMNF--GDW-----HLFRSVMLE 1220
Db 1490 GLSWTELARELNFSDVINQIRVENPNLSISQSFMLLKWVTRDGNKATTDALTSLVTKI 1549
QY 1221 RSVESQVPEDP-----RFLNENS--SAPVPHGESARRSSHTTELPTELSSQTP 1267
Db 1550 NRIDIVTLLEGPIDYGNISGTRSFADENNVDHVP-----DGH---PSFOVELETP 1598
QY 1268 YTNLNFSEELNTLGLDEGAPRHSNLSWQSQTRTTPSLSS-----LN 1308
Db 1599 MGLYWT--PNPFPQDD--HFSDISSIESPFTPSRLSDGLVPSQGNIEHPTGPPVVT 1653
QY 1309 SQSSSIEIKLTKV---QAEYRDAYREIVIAQMSQLEGTSSTISG-----RSP-- 1356
Db 1654 AEDTSLDSKMDSDSVTVTPADPLDVDESQKLCQSECAQCWASVPGIPNDGRQAEPLR 1713
QY 1357 HSYIYIGQSSGSIHSTLEQERKGEKLGKQEDGRKFLMKRGDVIDYSSSGVSTNEASP 1416
Db 1714 PQRKVMGSE-----QOEKRGSG-----P 1733
QY 1417 LDPITBEDSKDQSGSKLLPGKKSERPSLFTQDLKLGGLRYQKLPSDESGTGRVQ 1476
Db 1734 DEEVTEKVK-----SLFE-DIQLE--EVEAEEMTEPQGMALNRVQ 1772
QY 1477 ITPHCKMRTKRLKAKQREKASQPHESABPIRTFFIKAEYLSDDLKQSDSGVRSN 1536
Db 1773 -----RAELAMSSLAGWNETPSGLESQAARR-----LTGGLLDRLDSSDQARDS 1820
QY 1537 -----ESSPNHSLHNEAADDQLEKANLIELEDEGHSGKGM-----PHSL 1577
Db 1821 ITSYLTCGEPCKIANGNHT-----AEVPEAKAKPYFPESQNDIGKQSIKENLKPETHGC 1875
QY 1578 SGLQDPIIARMSTICSDKKSPSECS--LIASSPEESWPAQCKAYNLNRTPSTVTLLNNNTA 1635
Db 1876 GRTEEPV---SPLTAYQKSLETSKLVIEDAPKPCVPVGMKKM-----TRTT 1919
QY 1636 PTNRANONPDEIEG 1649
Db 1920 ADGKARLNLOEERG 1933

RESULT 9

T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PRT>
A;Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:9710549; PID:9710549; PIDN:AAB01604
A;Experimental source: strain C57BL/6J; kidney

QY 1599 SECS--LIASSPESWACOKAYNLRNTPSTVTLNNNTAPTNRANQNFDEIEG 1649
Db 1872 EETSKLVEDAPKPCVPGMKM-----TRTTADGKARLNQEBEG 1912

RESULT 10
T42713
ankyrin 3, splice form 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42713
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turto, J.
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: Z22237; MUID:95340633; PMID:7615634
A:Accession: T42713
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1943 <PEP>
A:Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:G710548; PID:G710550; PIDN:AAB01606
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 855/1
C:Function:
A:Description: supposed to play an important role in the polarized distribution of many
A:Note: major kidney ankyrin
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 6.2%; Score 551; DB 2; Length 1943;
Best Local Similarity 20.2%; Pred. No. 1.2e-23;
Matches 426; Conservative 289; Mismatches 666; Indels 724; Gaps 85;

QY 14 EENIPALKLECKDVERNECGQPLMAAEOQNVIEIKKXGANCNLEDLDNWT 73
Db 66 KEGHVEVSELLQREANVDAATKGNTHALHSLAGQAEVVKVLTNGANVAQSQNGFT 125

QY 74 ALISASKEGHIHVEELKSGASLEHRDGMGTALMACYKGRDVTVELLL----- 124
Db 126 PLYMAQENHLEVRFLLDNGASQSLATEDGFTPLAVALQOQHDQVVSLLLENDTKGVR 185

QY 125 -----SHGANPSVTGL----- 135
Db 186 LPALHIAARKDDTKAAALLLQNDNTADVESKSGFTPLHIAAHYCNINVTALLNRAAVID 245

QY 136 ---QYSVYPLIWAAGRHADIHLLQNGAKVNCSDKYGTPLVWAARKG----- 182
Db 246 FTARNDITPLHVASKRGNANVXLLLDGRGAKIDAKTRDGLTPLHCGARSQHEQVEMLLD 305

QY 183 -----HLECVKHLAMGADVDOEGANSMTALIVAVKGYTQS 219
Db 306 RSAPILSKTKNGLSPLHMAQGDHLCVQLLQHNVPDVTNDYLTALHVAACHGKYV 365

QY 220 VKBILKRNPNVNTDKDGTALMIASKE----- 247
Db 366 AKVLLDKASPNKALNGFTPLHIAACKNRIKRVWELLKHGASTQAVTESGLTPIHVAAF 425

QY 248 -GHIEIVODLLDAGTYNIPDRSGDTVLICAVGGHVEIVRALLQYADIDRQDNKTA 306
Db 426 MGHNVIVSQLMHGASPNNTNVRGETALHVAARSQAEVVRVYLDQGAQAEAKADQTP 485

QY 307 LYMAVEKGNATVRDILQCNPDTEICTKGETPLIKATKRNIEVVELLLDKGAKVSAD 366
Db 486 LHI SARLGKADI VQQLLQOQASPNAAATSGYTPPLHAAAREGHEDVAAFLLDHGASLSITT 545

QY 367 KGGDTPLHVAIRGRSRRLAELLR-----NPKDGRLL----- 398
Db 546 KKGFTPLHVAARYKLEVASLLLOKSASPDAAKSGSLTPLHVAHYDNDQKVALLLDQGA 605

QY 399 --YRPNKAGETPNIDC-SHOKSILTQI--FGAR-----HLSPTETDGMGLY 441
Db 606 SPAAAKNGTTPHIAAKNQMDIATSLLEYGADANAVTQGIASVHLAAQEGHVDVMSL 665

QY 442 DLYSSALADILSEPTMQP-----PICVGLYAOV 469
Db 666 LLSRNANVLSNKSGLTPLHAAQEDRVNVAEVLVNGQAHVDAQTKMGYTFPLHVGHCH--Y 723

QY 470 GSKG--SFLIK--KLEDEKMTFAGOOTPLFQFSWLVIVELTL-----LLCGGLGLV 516
Db 724 GNIKVNFLLQSHAKVNAKTNGYGTALHQAQOQGHTHIINVLLQNNASPNELTVNG--- 779

QY 517 FAFVDTNIAIATSLFLALIVIFFIVI----- 544
Db 780 -----NTALAIARGLYISVVDTLVKVTBEIMTTTTITEKHKNVPETMNEVLDMDSDEV 834

QY 545 -----YFGRGREGESNNWALSTRLARHIGYLELLFLKLMFVNPPPELPEQTT 591
Db 835 RKASAPEKLSDGHEYISDGEED-----AITGDTDKVLG-----PQDLKELGD 877

QY 592 KALPVRPLFTDYNRLSSVGGETSLAEMIATLSDACEREFGLATLRFVRPTESQKKK 651
Db 878 DSLPAE-----GYVG--FSLGARSASLR-----SFSSDRSYTLNRSSYAR--- 915

QY 652 WKTCCLPSPFVIFLFIIVGCIAGITILLAIFRVDP-KHLT-----VNAILISIASVWG 702
Db 916 -----DSMIEELLVPSKEQHLTTFREFDSLSRHSWAADLDNVLNVSFVHSGFL 968

QY 703 LAFVLCRTWQVLDLSLNSQ-----RK-----RLHSAASKLHLK-----KSE 740
Db 969 VSPMVDARG-----GSMRGSRRHGMRIIPPRKCTAPTRITCRVLKXHLKLANPPMVEGE 1023

QY 741 G-----FMKVLKCEVELMARMAKTIIDSFTQVQ-----TRL 770
Db 1024 GLASRLVEMPGAGQALGPVIVEIPHFGSMRGRKERELVLRSENGETKEHQFDSKNEDL 1083

QY 771 VVIIDGLDACEQDKVQLMDLTVRV--LFSKGPPIAFASDPHIIKAINQNLASVLRLDSN 828
Db 1084 AELLNGMD--EELDSPEELGTRKICRIIYK-DEPQYFA-----VVSRIKQESNQIGPEGG 1135

QY 829 INGHYMRNIVHLVFLVFNLSRGLSNARKFLV-TSATNGDITCSDTTGTQ-----EDTDRRV 882
Db 1136 I-----LSSITVPLVQASFPFGALTKRIRVGLQAQVPEETVKKI 1175

QY 883 SQNSLGMETKLSKTAIINRDTY--RRQWQRIITQMSFDLTLL--VTEWDFSDISP 937
Db 1176 -----LGNKATFSPITVVEPRRRKPKFET--MTIPVPPSGEGVNSYKGDATP 1223

QY 938 QTMRRLLNIVSVTG-----RLLRANQITFNWDLASWINLQWPTVTSWL-----I 984
Db 1224 NL--RL--CSITGTSPAQWEDITGTTPLTFIKDCVSTTNVSARF-----WLADCHQV 1274

QY 985 LYLEETEGLPDQMTLKT--YERISKNIPTTKOVEPLLEIDGDIRNF----- 1029
Db 1275 L---ETVGLASQLYRELICVPYMAKFWVFAKTNDP--VESSLRCFCMTDTRVDKLTLEQQ 1328

QY 1030 -----EYFLSRTPVLVARDVKFLPCTVNLDPKLR--ELIADVARE-----QINIGG 1077
Db 1329 ENFEVARSKDIEVLEGKPI--YVDCYGNLAPITKGQOQLVFNFYFKENRPFISKIRD 1386

QY 1078 LAYPPLP----LHEGPRPPSGYSQASVCSSASFNPGFPGVGVVSPQPHSSYISGLSGQP 1133
Db 1387 TSQEPGRLSFLKE--PKYTKGLPQTA--VCN-----LNITLPA 1421

QY 1134 HPFYNRAAVPATGSSLLL-----SSMTVDVCEK--LRQIEGLDONMPOYCTTIK 1182
Db 1422 HKAEKADRRQSPASALAKRYSYLTPESMSPQSCERTDIRMAIVAD-HLGLSWTELAR 1480

QY 1183 KANINGRVLSCQNIDELKEMANF--GDW-----HLFRSVLEMRSVESQVVP 1229
Db 1481 ELNFSVDEINQIRVENPNLSIOSFMLKKKVVTRDGNATTDALTSVLTINKINRIDIITLL 1540

QY 1230 EDP-----RFLNENS--SAPVPHGESARRSSHTLPTLTELSSQTPYTLNFSFEE 1276

Db 1541 EGPIFDYGNISGTRSFADENNVPDPV-----DGH-----PSQVLELETWGLYWT--P 1587
Qy 1277 LNTLGLDGGAPRHNLSQSQTRTPSLSS-----LNSQSSSIEIS 1317
Db 1588 PNPQQDD--HFSDISSIEFPRTPSRLSDGLVPSQGNIEHPGPPVPTAEDTSLSDS 1644
Qy 1318 KLTKV-----QAEYRDAYREYIAQMSQLEGGTSGSTISG-----RSSP--HSTVYIQGS 1365
Db 1645 KMDDSVTVTDPADPLVDVDESQKDLQSQEACQWASVFGIPNDGRQAEPPLRPQTRKVGMS 1704
Qy 1366 SSGSIHSTLQERQKQELQEDGRKFLMKRGVDIYSSGVSSTNDEASPLDPITEDE 1425
Db 1705 SE-----QOERKSGS-----PDEEVTEDEKV 1724
Qy 1426 KSDQSGKLLPGKSSBRPSLFOTDLKLGGLRYOKLPSPDEDSGTRGVQITHCSCMI 1485
Db 1725 K-----SLFE-DIQLE--EVEAEEMTEQGGQAMLRVQ-----RAEL 1758
Qy 1486 RTKRLKAKQRECAQPOEHSABPRTFIKAKELYSDALLDKKSDSGVRSN----- 1536
Db 1759 AMSSLAGQNETPSSGLESQAARR-----LTGGLDLRLDSSDQARDISITSLTCEP 1811
Qy 1537 ---ESSPNHSLHNEAADSQLEKANLIELEDEHSGKRGK-----PHSLSGLDPIIA 1586
Db 1812 GKIEANGHT-----AEVPEAKAKYPPEQSDNDIGKQSIKENLKPXTHGCGRTEEPV-- 1864
Qy 1587 RMSICSEDKSPSCS--LIASSPEESWPACQKAYNLNRTSTVTLNNTAPTNRANQNF 1644
Db 1865 --SPLTAYQKSLSETSKLVIEDAPKPCVPVGMKKM-----TRTTADGKARLNL 1910
Qy 1645 DEIEG 1649
Db 1911 QEEEG 1915

RESULT 11
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 09-Jul-2004
C;Accession: S37431; B39643; B39643; A40334; A49462; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: UNIPROT:Q01484; EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G4062
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302466; PMID:1830053
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OTL>
A;Cross-references: GB:X56957
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,3585-3924 <OTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009921; PMID:1833308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474, 'PE', 477-495 <TSE>

A;Cross-references: GB:W37123; NID:G178647; PIDN:AAA62828.1; PID:G178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-KD ankyrinB: structure of the major developmentally regulated domain and se
A;Reference number: A49462; MUID:94075409; PMID:8253844
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:G406287; PIDN:CAA81387.1; PID:G406288
C;Genetics:
A;Gene: GDB:ANK2
A;Cross-references: GDB:I27607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F;2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F;63-95/Domain: ankyrin repeat homology <AN01>
F;96-128/Domain: ankyrin repeat homology <AN02>
F;129-161/Domain: ankyrin repeat homology <AN03>
F;162-190/Domain: ankyrin repeat homology <AN04>
F;191-223/Domain: ankyrin repeat homology <AN05>
F;232-264/Domain: ankyrin repeat homology <AN06>
F;265-297/Domain: ankyrin repeat homology <AN07>
F;298-330/Domain: ankyrin repeat homology <AN08>
F;331-363/Domain: ankyrin repeat homology <AN09>
F;364-396/Domain: ankyrin repeat homology <AN10>
F;397-429/Domain: ankyrin repeat homology <AN11>
F;430-462/Domain: ankyrin repeat homology <AN12>
F;463-495/Domain: ankyrin repeat homology <AN13>
F;496-528/Domain: ankyrin repeat homology <AN14>
F;529-561/Domain: ankyrin repeat homology <AN15>
F;562-594/Domain: ankyrin repeat homology <AN16>
F;595-627/Domain: ankyrin repeat homology <AN17>
F;628-660/Domain: ankyrin repeat homology <AN18>
F;661-693/Domain: ankyrin repeat homology <AN19>
F;694-726/Domain: ankyrin repeat homology <AN20>
F;727-759/Domain: ankyrin repeat homology <AN21>
F;760-792/Domain: ankyrin repeat homology <AN22>
F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 550; DB 2; Length 3924;
Best Local Similarity 20.4%; Pred. No. 4.7e-23;
Matches 417; Conservative 279; Mismatches 727; Indels 626; Gaps 78;

Qy 14 EEENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKLLKNGANCNLEBDLWNT 73
Db 73 KEGHVLGVQELLGRGSSVDSATKKGNTALHIASLAGOAEVVKLVKEGANINAOQSNGFT 132
Qy 74 ALISASKEGHHIVVEILLKSGAS-----LEHRDMG--- 103
Db 133 PLYMAAOENHIDVVVKYLLENGANGQSTATGFTPLAVALQGHNOQVAILLNKTGKVR 192
Qy 104 -----GWTALMWACYKGRDVVVELL 123
Db 193 LPALHIAARKDDTKSAALLQNDHNADVQSGMVNRTESGFTPLHIAAHGVNVVATLL 252
Qy 124 LSHGANPSVTGLQSVYPIIWAAGRHADIVHLLQLONGAKVNCSDKYGTTPLVMAARKG- 182
Db 253 LNRGAADVDTA-RNGITPLHVASKRGNTNMVKLLLDLDRGGQIDAKTRDGLTPLHCAARSCH 311
Qy 183 -----HLECVKHLAMGADVDQEGANSNTALIV 210
Db 312 DQVVVELLLEGGAPLARTKNGLSPLHMAAQGDHVECVKHLQLQHKAPVDDVTLVDLTALHV 371
Qy 211 AVKGGYTSQVKEILKRNPNVNLTDKGNLTALMTASKS----- 247
Db 372 AAHCGHVTVTKLLDKEANPNARALNGFTPLHIAKKNRIKVMELLVKYGASIQAITESG 431
Qy 248 -----GHTEIVQDLLDAGTVYVNIPIRSQDVTILGAVRGGHVEIVRALLQKYADID 297
Db 432 LTPHVAAFMGHLNIVLLLLQNGASPDVTNIRGTALHMAARAGQVVEVRCLLRNGALVD 491

[illegible]

QY 1648 -EGIRETSQVILRPGSPNPNTA-----VQENLKSMAHKRSQSSYT 1688
Db 1751 QRGPPETGSP--RAGKEPSLWAPESAFSQEVOGDELQNPGEQVTBEQFT 1798

RESULT 13
A55575
N;Ankyrin 3, long splice form - human
N;Alternate names: ankyrin G
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 09-Jul-2004
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax
A;Reference number: A55575; MUID:95138209; PMID:7836469
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: UNIPROT:Q12955; GB:U13616; NID:G608024; PIDN:AAA64834.1; PID:G608025

C;Genetics:
A;Gene: GDB:ANK3
A;Cross-references: GDB:424503; OMIM:600465
A;Map position: 10q21-10q21
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: alternative splicing; peripheral membrane protein

F;73-105/Domain: ankyrin repeat homology <AN01>
F;106-138/Domain: ankyrin repeat homology <AN02>
F;139-171/Domain: ankyrin repeat homology <AN03>
F;172-200/Domain: ankyrin repeat homology <AN04>
F;201-233/Domain: ankyrin repeat homology <AN05>
F;234-266/Domain: ankyrin repeat homology <AN06>
F;267-299/Domain: ankyrin repeat homology <AN07>
F;300-332/Domain: ankyrin repeat homology <AN08>
F;333-365/Domain: ankyrin repeat homology <AN09>
F;366-398/Domain: ankyrin repeat homology <AN10>
F;399-431/Domain: ankyrin repeat homology <AN11>
F;432-464/Domain: ankyrin repeat homology <AN12>
F;465-497/Domain: ankyrin repeat homology <AN13>
F;498-530/Domain: ankyrin repeat homology <AN14>
F;531-563/Domain: ankyrin repeat homology <AN15>
F;564-596/Domain: ankyrin repeat homology <AN16>
F;597-629/Domain: ankyrin repeat homology <AN17>
F;630-662/Domain: ankyrin repeat homology <AN18>
F;663-695/Domain: ankyrin repeat homology <AN19>
F;696-728/Domain: ankyrin repeat homology <AN20>
F;729-761/Domain: ankyrin repeat homology <AN21>
F;762-794/Domain: ankyrin repeat homology <AN22>
F;795-827/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 526; DB 2; Length 4377;
Best Local Similarity 20.0%; Pred. No. 1.5e-21;
Matches 416; Conservative 266; Mismatches 664; Indels 734; Gaps 80;

QY 14 EEENIPALKLECKOVDNERECQYPLMAAQGVVEIVKELKKGANCLDLDNWT 73
Db 83 KEHVEVSELLQREANVDAATKNGTALHTASLAGQAEVVKVLVTNGANVNAQSQGFT 142
QY 74 ALISASKEGTHIVEELIKSGASLEHRDMGWTALMMACYKGRITDVVELL- 124
Db 143 PLYMAAQENHLEVVKFLDNGASQSLATEDGFTPLAVALQGHQDVVSLLENDTKGKVR 202
QY 125 -----SHGANPSVTGL----- 135
Db 203 LPALHIAARKDDTKAAALLQNDNNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVD 262
QY 136 ---QYSVVPIIWAAGRHADIVHLLLONGAKVNCSDKYGTTPLVWAARKG----- 182
Db 263 -FTARNDDITPLHVASKRGANVWKLLDRGAKIDAKTRDGLTFLHCGRSAGHEQVVELLD 322
QY 183 -----HLECVKHLAMGADVDEQANSMTALIVAVKGYTQS 219

Db 323 RAAPILSKTKNGSLPHLMATQGDHLNCVOLLQHNPVDDVTNDYLTALHVAACHYKV 382
QY 220 VBEILKRNPNVNLTKDNGNTALMIASKE----- 247
Db 383 AKVLLDKKANPNNAKALNGFTPLHIAKKNRIKIMWELLKHGASIOAVTESGUTPIHVAAP 442
QY 248 -GHIEIVQDLLDAGTVVNIPIPRSDGTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTA 306
Db 443 MGVNIVSQLMHGHGASPTNTVNRGETALHMAARSQAEVVRVIVQDGAQVEAKAKDDQP 502
QY 307 LYWAVEKGNATVVRDILQ--CNPD-----EICT 333
Db 503 LHISARLGRKADIVQQLQOGASPNATTSYGTPLHLSAREGHEDVAAFLLDHGASLITT 562
QY 334 KQGEPTLIKATKRNIEVVELLDCAKYSADVCKGDTPLHVAIRGRSRLAEILLRNPK 393
Db 563 KKGFTPLHVAAYKGLVANLLQKSASPDAGKSGLTPLHVAHYADNOKVALLL---D 619
QY 394 DGRLLYRPNKAGETPVNIDC-SHQKSILTQI--FGAR-----HLSPTETDGM 438
Db 620 QGASPHAAAKNGYTPHIAAKKNQMDIATLLEYGADANAVTRQGIASVHLAAQEGHVM 679
QY 439 LGYDLYSSALADILSEPTMQP-----PICVGLY 466
Db 680 VSLLLRNANVNLNKSGLTPLHLAAQEDRVNVAEVLVNOGARHVDATQTKGYTPLEVGCH 739
QY 467 AQWGSCK--SFLLK-KLEDEMTFAQQTEPFOFSWL--IVFLTLLCGGLGLVFAFP-- 520
Db 740 --YGNIKIVNVLQHSKAKVNAKTNGY--TPLHQAQOQGHTHIINVLQNN-----ASPN 790
QY 521 ---VDTNLAIAISLSPLALIYIFFIVYFGRRGESNMAWALSTRLARHIGYLELLPK 577
Db 791 ELTVNGTALGI-----ARRLGYISVDT 814
QY 578 LMPV-----NPPE-----LPEQTTKALPVRFLEFDYNRLLSV--GGET 613
Db 815 LKIVTEETMTTIVTEKHQNVPTWNEVLMSDDDEVKANAPEMLSGDEYISDVEGED 874
QY 614 SLA-----EMIAATLSDACEREBFGFLATL-----FRVFRTEESQKKKKWKTCC 658
Db 875 AMGTDDTKVLPQDLKELGDDSLPAGCYMGFSLCARSASLSRFSFSSRSTYLN--RSYAR 932
QY 659 PSFVIFLIVGCIAGITLLAIFRVPD--KHLT-----VNAILISIVSUGLAFVLC 709
Db 933 DSMWIELLVPSKEQHLTPTRFSDSLRHSWAADTLDNVNLVSSPIHSGFLVSEFMDA 992
QY 710 RTWQVLDLSLNSQRKLHSAASKLHKLKSEGPMKVL---KCEVELMARMAKTIDSTQ 765
Db 993 RG-----GSMRGSRHG-----MRILIPPRKCTAP-----TR 1019
QY 766 NQTRLVVIIDGLDACBQDKVLQMLDTRVLFSGPFIAPASDPHTIIKAINQNLASVLR 825
Db 1020 ITCRLV-----KHKLAN-----PPH-----GERRGISRLV 1047
QY 826 DSNINGHDYMRN-IVHLPVFLNSRGLSNARKFLVTSATNGDITCSDDTGTQEDTDRVSQ 884
Db 1048 EMGPAGAGLPVIVIEIPHFGSMRG--KERELIVLSRSENGE-----TWKEHQFD---SK 1096
QY 885 NSLGEMTKL--GSKTALNRDVTYRRQMORTITRQMSFDLTLLVTEDEWPSDISPQTMRR 942
Db 1097 NE--DUTELNGMDEELDSPEELGKKRICRIITKDF-----POYFAVVS--RIKQ 1142
QY 943 LLNIVSVTGRLLAN-----QITFNMDRLASWINLTEQWPFYRTSWLILYLEETEGLPDM 997
Db 1143 ESQIQGEGGILSTTVPLVQASFPPEGALTKRIRVGLQ-----AQPVDEI 1188
QY 998 TLKTM-----YERISKNP-----TTKQVEPLLEI 1022
Db 1189 VKKILGNKATFSPIVTVPEPRRRKPHFTWTVPPPPSGEGVNGYKGDFTPNRLLC 1248
QY 1023 DG-----DIRNFEVFLSSRTPVLVARDVKTFPLCTVNLDPKLRBIIADVRAARQIN 1074

Db 1249 TGGTSPAQWEDI-----TGTPLTFTFKDCVSP-----TTNV--SARFWLADCHQVLETVG 1296
Qy 1075 IGLGAYPLPLHEGPPRPSGYSPASVCSASFNGPPGGVVSQPQSHSYSGLSGPOH 1134
Db 1297 LATQLYREL-----ICV-----PYAKV-----VFAKNDP-- 1323
Qy 1135 PPYNRAAVPATGSSLLSMTVDVCEKLRQIEGLDQNMPOYCTTIKANINGRVLQ 1194
Db 1324 -----VESSRLRCFCMTDDKVDKLEQOE----- 1346
Qy 1195 NIDELKEMAMNFGDHLFRSNVLEMRSVESOVVDPEDRFLN-ENSSAPVPHGE-----S 1248
Db 1347 NPEEV-----ARSDIEVLEGKPIYDVCYCNLAPLTKGGQOOLVFN 1386
Qy 1249 ARRSHTTELP-----LTLESSTOPYTLNFSFELNLTGLDEGAPRNSLWSOQTRTPSL 1304
Db 1387 FYSFKENRLPSIKIRDTISOECGRSLFKERKTKGLPQAVCNLNTLPAKKETES- 1445
Qy 1305 SLSNQSSSIEISKLTDKQVAYRDYRE---YIAQMSOLEGGTSGSTISGRSSPHSTYY 1361
Db 1446 -----DQDDEIE---KTRRQSPASLAKRYSYLTPEGMIERSTGAT---RSLP-TTV- 1492
Qy 1362 IQSSSGSGSIHLEQERKEGELQEDGRKSFMLKRGDVIDYSSGVSTNEASPLDIT 1421
Db 1493 -----SYKPFSTRPYQSWTTAPITVGPAGS-----GFTSLSSSSNTPSASPLKSIW 1541
Qy 1422 BEDEKSDQSGSKLPGKKSERPSLFQTDLKLKGGGLRYQKLPSEDESGTGRVOITPHC 1481
Db 1542 -----SVTSPSIKSTL-----GASTTSSVKSISDV 1567
Qy 1482 SKMIRTKLAKQRECAPQEHASPIRTFKAEY-----LSDALLDK-----KDS 1528
Db 1568 ASPIRSLRT-----MSSPIKTVWSQSPYNIQVSSGTLARAPAVTEATPLKGL 1614
Qy 1529 SDSGVRSNESPNSHNEADDLSQLEKANLIELEDEHSGKRGMPHS----- 1576
Db 1615 ASNSTFSRTSP-----VTTAGSLERSSTMTPTPASPKSNINMYSSLPFKSIITSA 1668
Qy 1577 -----LSGLQDPIIARMSICSEDKSPSECSLIASSPEESWPA----- 1614
Db 1669 PLISSPLKSVVSPKVRVDVISAKITWASS-----LSSPVKOMPQHAVALVNGSISPLKY 1725
Qy 1615 COKAYNLNRTPSTVTLNNN-TAPTNRANQ-----NFEIEGIRETSQVI-----LRPGSPN 1665
Db 1726 ASSSTLINGCKATATLOEKISSATNSVSVSAATDTVEKFVSTTTAMPFSPRLSYVSA 1785
Qy 1666 PTAQVENLKSMAHQRSSSYTRLSKDASELHAASSEST 1705
Db 1786 PSFQSLRTPSAS-----ALYTSL-----GSSISATTSVT 1816

RESULT 14
S37771
ankyrin, erythrocyte - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S37771
R:Birkenmeter, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an
A:Reference number: S37771; MUID:93252825; PMID:8486643
A:Accession: S37771
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1848 <NR>
A:Cross-references: UNIPROT:061302; EMBL:X69063; NID:g311816; PIDN:CAA48801.i; PID:g3118
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:48-80/Domain: ankyrin repeat homology <AN01>
F:81-113/Domain: ankyrin repeat homology <AN02>
F:114-146/Domain: ankyrin repeat homology <AN03>
F:147-175/Domain: ankyrin repeat homology <AN04>
F:176-208/Domain: ankyrin repeat homology <AN05>

F:209-241/Domain: ankyrin repeat homology <AN06>
F:242-274/Domain: ankyrin repeat homology <AN07>
F:275-307/Domain: ankyrin repeat homology <AN08>
F:308-340/Domain: ankyrin repeat homology <AN09>
F:341-373/Domain: ankyrin repeat homology <AN10>
F:374-406/Domain: ankyrin repeat homology <AN11>
F:407-439/Domain: ankyrin repeat homology <AN12>
F:440-472/Domain: ankyrin repeat homology <AN13>
F:473-505/Domain: ankyrin repeat homology <AN14>
F:506-538/Domain: ankyrin repeat homology <AN15>
F:539-571/Domain: ankyrin repeat homology <AN16>
F:572-604/Domain: ankyrin repeat homology <AN17>
F:605-637/Domain: ankyrin repeat homology <AN18>
F:638-670/Domain: ankyrin repeat homology <AN19>
F:671-703/Domain: ankyrin repeat homology <AN20>
F:704-736/Domain: ankyrin repeat homology <AN21>
F:737-769/Domain: ankyrin repeat homology <AN22>
F:770-802/Domain: ankyrin repeat homology <AN23>

Query Match 5.8%; Score 516.5; DB 2; Length 1848;
Best Local Similarity 19.6%; Pred. No. 1.2e-21;
Matches 397; Conservative 281; Mismatches 651; Indels 697; Gaps 81;
Qy 21 LKALLEKCKYDERNECQOTPLMLAABQGNVEIVKELKNGANCNLELDNWTALISAK 80
Db 98 VRELNVYGANVAQSQKGFPLYMAAQENHLEVKVFLLENGANQNVATEDGFTPLAVALQ 157
Qy 81 EGH----- 83
Db 158 QGHENVVAHLINYGKGVKRLPALHIAARNDDTRTAVALQLQNDPNPVLSTKGTFTPLHIA 217
Qy 84 -----IHVEILLKSGASLEHRDMGWTALMWACVKGRTDVVELLSHGANSVTLGLOYS 138
Db 218 AHYENLVNAQLLNRGASVNFPPQGITPLHIAISRGVNIWVRLLDLDRGAQIE-TRTKDE 276
Qy 139 VYPIIWAAGRHDIVHLLQNGAKVNGSDKYGTTPLVMAARKHLEKVCVGHLLAMGADV 198
Db 277 LTPHCAARNGHVRISIELLDHGAPIQAKTKNGLSPIHMAAQGHLDVCRLLQYNAEID 336
Qy 199 -----QEGA-----NSMTALIVAVKGGYTSQSKELK 225
Db 337 DITDLHTPLHVAACHGHRVAKVLLDKGAKPNSRALNGFTPLHIAKQNHIRVMLELLK 396
Qy 226 RNPVNLTKDGNLTALMASKEGHEIIVODLLDAGTYVNIPIRSGDTVLIGAVRGHVEI 285
Db 397 TGAISIDAVTESGLTPLHVASFMGHLPIVKNULQKASPNVSNVKTETPLHVAARAGHTEV 456
Qy 286 VRALLQKYADIDIRGQDNKTALYWAVEKGNATMVRDILQCNPDTEICTKOGETPLIKATK 345
Db 457 AKYLLQNKAKANAKAKDDQTPHCAARIGHTGWVKLLLENGASPNLATTAGHTPLHTAAR 516
Qy 346 MRNIEVELLLDKGAKVSAVDKKGDTPLHVAIRGRSRLLAELLENPKDGRLLYRPNKAG 405
Db 517 EGHVDTLALLEKEASQACMTKGTFTPLHVAARYKVKRLAEELLEHDA-----HPNAA 570
Qy 406 E---TPYNIIDCSHQ-----KSIL-----TOIFGAR----- 427
Db 571 KNGITPLHVAHVHNNLIVKLLLPGRGSPHSPANNGVYTPHIAAKQNOIEVARSLQYGG 630
Qy 428 -----HLSPTETDGMGLYDLYSSALADILSEPTMOP-----PIC-- 462
Db 631 SANAESVQGVTPPLHAAQEGHTEMVALLSKQANGNLGNKSGLTPLHLVLSQEGHVPVADV 690
Qy 463 -----VGLYAQWGSK--SFLIKKLED-EMKTFACQQTPEPLQFQSWL 501
Db 691 LIKHGVTVDAITRMGMVTPHVAISHYGNIKLVKFLQHQADYNAKTKLGY--SPLHQAAQQ 748
Qy 502 --IVFLTLLCGGLGVFAFPVD-----TNLATAISLPLALIYIPFIVYFGRRGE 553
Db 749 GHTDIVTLLKXG-----ASPNEVSSNGTTPPLAKRGLYISVTDVLKVT----- 794
Qy 554 SNNWAWALSTRLARHIGYLELLFLKMFVNPPELPEQTTKALPV---RPLFTDYNRLSS-- 608

795 --DTSVVLVDKRMSPYPTVDEILDVS-----EDEGTAHISIMGDELVGSKAERRDSRD 848
609 VGETSLAEMIATLSDACEREFGLATRLPR-----VFRTB-ESQKKKWKTCCL 658
849 VGEKELLDPVKLDQVVE-----SPAIPRIPCTVPETVIRSEDOEQASKEYDEDSLI 902
659 PSFVIFLIVGCIAGITLAIIFRVDPKHLTVNAILISIASVGLAFVLCNCRTWQVLDLS 718
903 PS-----SPATETSONI-SPVASPVHTGFLVSF-----WVDA 933
719 LINSORKRLHSAASKLHLKXSEGMVKLKCEVELMARMMAKTIDSFTQNTQRLVVIIDGLD 778
934 RCGSMGRSRHN-----GLRVVIP-----PRTCAAPRITCLV----- 966
779 ACSQDKVQLMDTVRVLFSGKPFIAIPASDPHIIIIKAINONLSVLRDSNINGHDYMRN- 837
967 -----KPKQLNTPPL-----AEEGLASRIIATLGPPTGAQFLSPV 1001
838 IVHLPVPL-NSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSK 896
1002 IVEIPFASHGRG--DRELVLRSENGSV-----WKEHKSRYGESYLDQILN-GMD 1049
897 TALNRRTYRRRQMTITRQMSFDLTKLIVT---EDWFSDISPQTMRLNINVSVTGRL 953
1050 EELGSLEELEKRVCRIT--TDPLFYVIMSLRQCQ-YDTIGPEG-----GSLRSKL 1099
954 LRANOITFNWDRLASINLTQWPYRTSWILLYLETEGLPDQWTLKMTYERISKNIPTT 1013
1100 VPLVQATFPFENAVTNVKLALQ-----AQVPDELVTKLGNQ----- 1137
1014 KDVEPLLEIDGDIRNFEVFLSSRTPLVARDVKTFPLCTVNLDPKLEIIVADVRAAREQI 1073
1138 -----ATFSP-IVTVPRRRKFRPI----- 1157
1074 NIGSLAYPPLPLEHGGPPRPSGYQPAS---VCSSASFNPGFPFPGVVPSPHSSYSGLS 1130
1158 --GLRIPLPSPWTDNPR-DSGEGDTTSLRLCSVI-----GGTDQAQ-----WEDIT 1201
1131 GPQHPFY-NRAA-----VPAIG-----SSILLSSMTVDVUCEKLRIQIEGLD 1170
1202 GTTKLIYANECANFTNVSARFWSLDCPRTAEAVHFATLLYKELTA----- 1247
1171 QNMWPOYCTTIKKANIN---GRVLSQCNIDE-LKEMAMNFGDWHLFRSMVLEM----- 1220
1248 --VPYMAKFIPIFAKMDAREGLRCYCMTDDKVDKTLBOHENFVEVARSRLIEVLEGP 1304
1221 --RSVESQVVP-----EDPRFLNENSSAPVPHGESARR-----S 1252
1305 LFAELSGNLVPVKAAQOGRGFHOSFRENRLAIPVKVRDSSREPFGFLSLFKTKMYEDT 1364
1253 SH-----TELPTELTS-----SOTPYTLNFSFEELATLGLDEGAPR----- 1288
1365 QHILCHLNITMPPCTKGSGAEDRRRTTLPULTRYLSILSESLRGLTSDTDVEMRMAVIRE 1424
1289 HSNLWSQSTR-----RTPSLSLNSQDSSI-----EISKLTDKVQAEY 1327
1425 HLGSLWELARELOFSVEDINRIEVENPNSLLDSTALLTIWVDREGENAKM-ENLYTAL 1483
1328 RDAYREYIAQMSLEGGTGSTISGRSPHSTYYIGOSSSGGSIHSTLEOERGKEGELKQ 1387
1484 RNIDRSEIVNMLE-----VSGRQS-----RNLKPER----- 1509
1388 EDGRKSLMKRGDVIDYSS-----SGVSTNEASPLDPITEDEKSPQSGSKLLPGKSS 1441
1510 RHGREYSLSPSQVNGYSSLODELLSPASLOYP-SPLCADQVWNEVTVDAIPLAATE 1568
1442 ERPSLFOTDLKLGGLRYQKLPSEDESGTGRVQITPHCSKMIRTKRLKAKORECASPO 1501
1569 HDTMLENSDMQVMSAGLT-PSLVTAEDSS-----LECSKAE 1603
1502 EHSAPERTFTKAEYLSALLDKKSDSDSVRSNENSSPHNSLHNEAADSQLEKANILE 1561
1604 D-----SDAIPWK-----LEGASDPTQPGGSDQLVE 1633

QY 1562 LEDEGHSKRGMPHSISGLQDPIIARMSICSSEDKKSPSECSLI-----A 1605
Db 1634 DTVDSDATNGLADLLG--QORVHAKIT-----DSPSVQVLDRSQARTLWDKOGSTA 1685
QY 1606 SSPEE-----SW--PACQAKAYNLNRPSTPVTNNNTAPTNRANQNFDEI-----EGI 1650
Db 1686 VHPQEAQTQSSQBEVTOGPHSFQRRITTI-----QGPEPGALQEVYQVLVSTREHVQGRP 1740
QY 1651 RETSQVILRPGSPNPNTA-----VQENELKSMHAKRSQSSSYT 1688
Db 1741 PETGSP--KAGKEPSLWAPESAFSQEVOGDELQNPGEQVTBEQFT 1784
RESULT 15
T42714
ankyrin 3, splice form 2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T42714
R/Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A/Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A/Reference number: Z22237; MUID:95340633; PMID:7615634
A/Accession: T42714
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1765 <P>
A/Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605
A/Experimental source: strain C57BL/6J; kidney
C/Genetics:
A/Gene: Ank3
A/Map position: 10
A/Introns: 1387/1
C/Superfamily: ankyrin; ankyrin repeat homology
C/Keywords: alternative splicing

Query Match 5.8%; Score 510.5; DB 2; Length 1765;
Best Local Similarity 20.1%; Pred. No. 2.6e-21;
Matches 396; Conservative 278; Mismatches 641; Indels 653; Gaps 78;
QY 14 EENIPALKALLEKCKVDNERNECGQTPMLAAEQGNVEIVKELKNGANGCNLEDLQNW 73
Db 66 KEGHVEVSELLQREANVDAATKKGNTALHIASLAGOAEVVKVLTNGANVNAQSQNGFT 125
QY 74 ALISASKEGHIHVEELLKSGASLEHRDMCGWTALMWACVKGRTDVVELL----- 124
Db 126 PLTMAAQENHLEVYRFLDNGASQSLATEDGFTFLAVALQQQHDQVVSLELNDTKGVR 185
QY 125 -----SHGANPSVTGL----- 135
Db 186 LPAIHIAARKDDTKAAALLLQNDTNADVRSKSGFTPLHTAAHYGNINVATLLNRAA 245
QY 136 ---QYSYPIIWAAGRHADIHLLLONGAKVNCSDKYGTGTPPLVWAARK----- 182
Db 246 FTARNDDITPLHVASKEGNANMVKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVEMLLD 305
QY 183 -----HLSCVKHLLAMGADVDEGANSMTALIVAKGVGYTQS 219
Db 306 RSAPILSKTKNGLSPLHMAATQGDHLNCVQLLQHNVPVDDVTNDYLTALHVAACHG 365
QY 220 VKELKRNPNVNITDKGNTALMIASKE----- 247
Db 366 AKVLLDKKASPNKALNGFTPLHIAACKKNRIRVWELLKHGASIQAVTESGLTPIHVA 425
QY 248 -GHIEIVQDILLDAGTVNIPDRSGDVTVLICAVGGHVEIVRALLQKVADIDIRQDNKTA 306
Db 426 MGHVNIQSQLMHHGASPNNTNVRGETALHMAARSQAEEVVRVYLVQDGAQVEAKKDDPT 485
QY 307 LYVAVEKGNATWVDILQCNPDTEICTKGETPLIKATKMRNIEVVVELLLDKGAKVSAVD 366
Db 486 LHISARLGRADIVQQLLQOQASPNAAATSGYTPPLHIAAREGHEDVAFLDHGASLITT 545

QY	367	KKGDTPLHVAIRGRSRLAEALLR- : : : : : :	-----NPKDGRLL-----	398
Db	546	KKGFTPLHVAAYKGLVASLLQKSAPDAAGSKGLTPLHVAAYHQNQKVALLLDQGA	605	
QY	399	--YRPNKAGETPYNIDC--SHQKSILTQI--FCAR-- : : : :	-----HLSPTETDGMGLY 441	
Db	606	SPHAAAKNGYTPHIAAKGNQMDIATSLLEYGADANAVTROGIIASVHLAAQEGHVDVWSL	665	
QY	442	DLYSSALADILSEPTMOP- : : : :	-----PICVGLYAQM 469	
Db	666	LLSNANVNLSNKSGLTPLHLAAQEDRVVAEVLVNOGAHVDAQTOMGYTFLRVGCH--Y	723	
QY	470	GSRK--SFLLK--KLEDEMYTPAQQTEPIFQPSWLIIVFTL- : : : :	-----LLCGGLGV 516	
Db	724	GNIKIVNFLQHSKAVNAKTNGYTAHQAAQOQGHTHIINVLQNNASPNELTVNG--	779	
QY	517	FAPFVDVTLAIALISLFLAIYIIFIVI- : : : :	-----544	
Db	780	-----NTALAIARRLGYISVVDTLKVYTEIIMTTTTTEKHKNVPETMNEVLDMSDEV	834	
QY	545	-----YFGRREGESNNAWALSTRLARHIGYLELLFKL- : : : :	-----MFVN 582	
Db	835	RKASAPKLSGEYISDGEEDKCTWFKIPKVQ- : : : :	-----EVLVKSEDAITGDTDKVLG 886	
QY	583	PPLEPQOTKALPVRFELTDYNRSLSSVGGETSLAEMIATLSDACEREFGLATLFRVFR	642	
Db	887	PODUKEGGDLSLPAE- : : : :	-----SPSSDRSXYTLNR 928	
QY	643	TEESQGGKKWKTCCLPSFIFLPIVGCIIAGITLLAIFRVPD-KHLT- : : : :	-----VNAI 693	
Db	929	SSYAR- : : : :	-----DWMIEELLVSKQHLTFREFDSDSLRYHWSWAADTLDNVNLV 977	
QY	694	LISIASVVGIAFVLCNTWQVLDLSLNSQ- : : : :	-----RK-----RLHSAASKLHL- 737	
Db	978	SSPVHSGFLVSFMVDARG- : : : :	-----GSMGRSHHGMRIIIPPRKCTAPTTRICRLVRRHKL 1032	
QY	738	-----KSEG- : : : :	-----FMKVLKCEVELMARMAKTIDSTQN 766	
Db	1033	NPPMWEGEGLASRLVEMGPAGAQLGPVIVEIPHGSMGKRELIIVLSENGETWKEH	1092	
QY	767	Q-----TRLVVIIDGLDACEQDKVQLMDLTVRV--LFSKGPFTAIFASDPHIIIKAINQN	819	
Db	1093	QFDSKNEDLAELLGMD--EELDSPEELGKTRICRIITK-DFQYFA-----VVSRIKQE	1144	
QY	820	LNSVLRDSNINHGYMRNIVHLPVFLNSRGLSNARKFLV-TSATNGDITCSDTTGTQ--	875	
Db	1145	SNQIGPBGGI- : : : :	-----LSSTVPLVQASPEGALTKRIRVGLQAQP 1184	
QY	876	--EDTDERVSONSIGEMTKLGSKTALNRDTY--RREQMORTITRONSPDLTKLL-----VT	928	
Db	1185	VPETVKKI- : : : :	-----LGNKATFPIVTVPRRRKFKHPIT--MTIIVPPPPSGEVS 1232	
QY	929	EDWFSDISPOTMRLNLNIVVTG- : : : :	-----RLLRANQITFNMDRLASNLNLTEQMPYRT 980	
Db	1233	NGYKGDATPNL--RL--CSITGTSQAWEIDTGTPLFIKDCVSFTTVNSARF-----	1284	
QY	981	SWL-----ILYLBETEGLPDMTLKTYERISKNIPTTKDVEPLLEIDGDIRNEFVLSS	1035	
Db	1285	WLADCHQVL--ETVGLASQLYRELIC-----VPY- : : : :	-----MAKFWVFAKT 1321	
QY	1036	RTPLVARDVKTFLPCTVNLDPKLEIADVRAREQINIGGLA-YPPLPLHSGPPRPPS	1094	
Db	1322	NDP-----VESSLRCFCWTDDRVDKTL- : : : :	-----EQEENFEVARSKDIEVLGKPIYVD 1369	
QY	1095	GYSQPASCVCSSAFNGFPFGGVSPQPHSSYYSGLSGPQHFF-----YNRAAVPATGSSLL	1450	
Db	1370	CYGNLAPLTK-----GG-----QQLVNFYS-FKENRLPFSIKIRDTSQBPGRLSFL	1416	
QY	1151	LSMTVDVCEKURQIEGLDQNMMPQYCTTIKANINGRVLSCQNDIDELKEMAMFPGDW	1210	
Db	1417	KEPKTT-----KGLPQTAV-----CNMLITLPAHKKAERDROS-----	1451	

Search completed: February 5, 2005, 19:27:51
Job time : 58 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:17:40 ; Search time 188.5 Seconds
(without alignments)
4658.970 Million cell updates/sec

Title: US-10-021-571-2
Perfect score: 8853
Sequence: 1 MSVLISQSVINYVEENIPA.....ELHAASSTGFGHERESIL 1715

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8853	100.0	1715	2 Q9ERD4	Q9erd4 rattus norv
2	8639.5	97.6	1762	2 Q9EQG6	Q9eqg6 rattus norv
3	8212	92.8	1777	2 Q9ULH0	Q9ulh0 homo sapien
4	7892	89.1	1693	2 Q80TG7	Q80tg7 mus musculu
5	7693	86.9	1554	2 Q80SX9	Q80sx9 mus musculu
6	6233.5	70.4	1680	2 Q7T163	Q7t163 brachydanio
7	6219.5	70.3	1672	2 Q6P7Y1	Q6p7y1 brachydanio
8	5073.5	57.3	1031	2 Q9UF42	Q9uf42 homo sapien
9	3097	35.0	691	2 Q7Z322	Q7z322 homo sapien
10	2732	30.9	664	2 Q6MZU2	Q6mzu2 homo sapien
11	2421.5	27.4	1604	2 Q7KVPE	Q7kvp5 drosophila
12	2421.5	27.4	1626	2 Q7KVPE	Q7kvp6 drosophila
13	2421.5	27.4	1678	2 Q9W210	Q9w210 drosophila
14	2408	27.2	543	2 Q9H9E4	Q9h9e4 homo sapien
15	2329.5	26.3	1354	2 Q7PZV2	Q7pzy2 anopheles g
16	1876	21.2	1089	2 Q7YU92	Q7yu92 drosophila
17	1848.5	20.9	1398	2 Q20109	Q20109 caenorhabdi
18	1097	12.4	239	2 Q8K0F2	Q8k0p2 mus musculu
19	1024	11.6	246	2 Q9NT37	Q9nt37 homo sapien
20	706	8.0	257	2 Q7OMF9	Q7gmf9 anopheles g
21	599.5	6.8	129	2 Q9H899	Q9h899 homo sapien
22	556	6.3	1880	1 ANK1_HUMAN	P46157 homo sapien
23	555	6.3	426	2 Q8YTG9	Q8ytg9 anabaena sp
24	555	6.3	1856	2 Q99407	Q99407 homo sapien
25	551	6.2	1943	2 Q61307	Q61307 mus musculu
26	550	6.2	3924	1 ANK2_HUMAN	Q01484 homo sapien
27	545	6.2	1719	2 Q13768	Q13768 homo sapien
28	543	6.1	1887	2 Q23364	Q2z364 homo sapien
29	531.5	6.0	1280	2 Q6N064	Q6n064 homo sapien
30	531.5	6.0	1863	2 Q7Z315	Q7z315 homo sapien
31	529	6.0	2622	2 Q70511	Q70511 rattus norv

32	527.5	6.0	1862	1 ANK1_MOUSE	Q02357 mus musculu
33	526	5.9	4377	1 ANK3_HUMAN	Q12955 homo sapien
34	523.5	5.9	1375	2 Q6ZSE6	Q6zse6 homo sapien
35	516.5	5.8	1848	2 Q61302	Q61302 mus musculu
36	510	5.8	1726	2 Q8VC68	Q8vc68 mus musculu
37	509.5	5.8	2206	2 Q7QAV8	Q7qav8 anopheles g
38	505.5	5.7	1136	2 Q9N180	Q9n180 bos taurus
39	504.5	5.7	1219	2 Q8C8R3	Q8c8r3 mus musculu
40	503.5	5.7	843	2 P97582	P97582 rattus norv
41	500.5	5.7	1486	2 Q8TEF1	Q8tef1 homo sapien
42	499.5	5.6	1159	2 Q9NCF8	Q9ncp8 drosophila
43	499.5	5.6	1571	2 Q7KU92	Q7ku92 drosophila
44	498	5.6	792	2 Q7Q172	Q7q172 anopheles g
45	493.5	5.6	830	2 Q96I86	Q96i86 homo sapien

ALIGNMENTS

RESULT 1
Q9ERD4
ID Q9ERD4 PRELIMINARY; PRT; 1715 AA.
AC Q9ERD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ankyrin repeat-rich membrane-spanning protein.
GN Name=ARMS;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20585245; PubMed=1150334;
RA Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;
RT "An evolutionarily conserved transmembrane protein that is a novel
RT downstream target of neurotrophin and ephrin receptors.";
RL J. Neurosci. 21:176-185(2001).
DR EMBL; AF313464; AAC34167.1; -.
DR HSSP; P09959; 1SW6.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 1715 AA; 190532 MW; 505593C4A19A6CDC CRC64;

Query Match	100.0%;	Score	8853;	DB	2;	Length	1715;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1715;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MSVLISQSVINYVEENIPALKALLERCKVDVERNECGQTPLMLAAEQGNVEIVKELKN	60				
Db	1	MSVLISQSVINYVEENIPALKALLERCKVDVERNECGQTPLMLAAEQGNVEIVKELKN	60				
Qy	61	GANCNLEDLNDWTALISASKEGHIHIVEELIKSGASLEHRDMGWTALMWACYKGRD	120				
Db	61	GANCNLEDLNDWTALISASKEGHIHIVEELIKSGASLEHRDMGWTALMWACYKGRD	120				
Qy	121	ELLISHGANPSVTGLQYSVPIIWAAGRHADIVHLLQNGAKVNCSDKYTTPLVWAAR	180				
Db	121	ELLISHGANPSVTGLQYSVPIIWAAGRHADIVHLLQNGAKVNCSDKYTTPLVWAAR	180				
Qy	181	KGHLECVKHLAMGADVDOEGANSMTALIIVAVKGGYTQSVKILKRNPNVNLTKDGN	240				
Db	181	KGHLECVKHLAMGADVDOEGANSMTALIIVAVKGGYTQSVKILKRNPNVNLTKDGN	240				
Qy	241	LMTASKEGHIIEVDLLDAGTYVNIIDPSGDTVLIGAVRGHVEIVRALLQKYADIDIR	300				
Db	241	LMTASKEGHIIEVDLLDAGTYVNIIDPSGDTVLIGAVRGHVEIVRALLQKYADIDIR	300				

Db 241 LMIASKEGHIEIYQDLDLAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Qy 301 QDNKTALYWAVEKGNATWRDILQCNPDTEICTKDGETPLIKATKRNIEVVELLDKGA 360
Db 301 QDNKTALYWAVEKGNATWRDILQCNPDTEICTKDGETPLIKATKRNIEVVELLDKGA 360
Qy 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGLLRPNKAGETPNYIDCSHOKSIL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRLAELLRNPKDGLLRPNKAGETPNYIDCSHOKSIL 420
Qy 421 TQIFGARHLSPETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKX 480
Db 421 TQIFGARHLSPETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKX 480
Qy 481 EDEMKTFAQOQTEPLFOFQSWLIIVFLTLCCGGLGLVAFPPVDNTLAIASLSFLALIYIF 540
Db 481 EDEMKTFAQOQTEPLFOFQSWLIIVFLTLCCGGLGLVAFPPVDNTLAIASLSFLALIYIF 540
Qy 541 FIVIFGGRREGSWMWALSTRLARHIGYLLPKLMFVNPPPELPEQTTKALPVRELF 600
Db 541 FIVIFGGRREGSWMWALSTRLARHIGYLLPKLMFVNPPPELPEQTTKALPVRELF 600
Qy 601 TDYNRLLSSVGGTSLAEMIATLSDACERERGFATRLFRVFRTEESQGGKKWKTCCLPS 660
Db 601 TDYNRLLSSVGGTSLAEMIATLSDACERERGFATRLFRVFRTEESQGGKKWKTCCLPS 660
Qy 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLCNTWQVLDLSLL 720
Db 661 FVIFLFIIVGCIAGITLLAIFRVDPKHLTVNAILISIASVVGGLAFVLCNTWQVLDLSLL 720
Qy 721 NSORKRLHSAASKLHKLSGFGMKVLKCEVELMARAKTIDSTFQNTQRLVWIIDGLDAC 780
Db 721 NSORKRLHSAASKLHKLSGFGMKVLKCEVELMARAKTIDSTFQNTQRLVWIIDGLDAC 780
Qy 781 EQDKVQLMDLTVRVLSKGFPIAFASDPHIIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Db 781 EQDKVQLMDLTVRVLSKGFPIAFASDPHIIIIKAINQNLNSVLRDSNINGHDYMRNIVH 840
Qy 841 LPVFLNSRGLSNARKELVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKELVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 900
Qy 901 RRDYRRRQWQRIITRQMSFDLTKLAVTEWDFSDISQPMWRLNLNVSVTGRLLRANQIT 960
Db 901 RRDYRRRQWQRIITRQMSFDLTKLAVTEWDFSDISQPMWRLNLNVSVTGRLLRANQIT 960
Qy 961 FNWDLASWINLQWPIRTSWLILYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPLL 1020
Db 961 FNWDLASWINLQWPIRTSWLILYLEETEGPLDQMTLKTMYERISKNIPTTKDVEPLL 1020
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQINIGGLAY 1080
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFPLCTVNLDPKLEIADVRAAREQINIGGLAY 1080
Qy 1081 PPLPLHEGPPRPSGVSQPSASVCSASFNGPFGVVSOPHSYTYGLSGPOHPFNRA 1140
Db 1081 PPLPLHEGPPRPSGVSQPSASVCSASFNGPFGVVSOPHSYTYGLSGPOHPFNRA 1140
Qy 1141 AVPATGSSLLSSMTVDVCEKLRQIBGLDQNMWQYCTTIKKANINGRVLSCNIDELK 1200
Db 1141 AVPATGSSLLSSMTVDVCEKLRQIBGLDQNMWQYCTTIKKANINGRVLSCNIDELK 1200
Qy 1201 KEMAMNFGDWHLFRSMVLENRVSVESQVVPDPRFLNENSSAPVPHGESARRSHTEPLT 1260
Db 1201 KEMAMNFGDWHLFRSMVLENRVSVESQVVPDPRFLNENSSAPVPHGESARRSHTEPLT 1260
Qy 1261 ELSSTQPYTLNFPSEELNTLGLDEGAPRHSNLWSQSTRTPSLSSNSQDSIEISKLT 1320
Db 1261 ELSSTQPYTLNFPSEELNTLGLDEGAPRHSNLWSQSTRTPSLSSNSQDSIEISKLT 1320
Qy 1321 DKVQAEYRDYREYIAQMSOLEGTTGSGSTISGRSSPHSTYYIQQSSGGSIHSTLEOERG 1380
Db 1321 DKVQAEYRDYREYIAQMSOLEGTTGSGSTISGRSSPHSTYYIQQSSGGSIHSTLEOERG 1380

Qy 1381 KEGELKQEDGRKSFMLMRGDVIDYSSSGYSTNEASPLDPTTEDEKSDOSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGRKSFMLMRGDVIDYSSSGYSTNEASPLDPTTEDEKSDOSGSKLLPGKKS 1440
Qy 1441 SERPSLFQTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHCSQWIRTKLKAQORECASP 1500
Db 1441 SERPSLFQTDLKLKGGGLRYQKLPSDEDESGTGRVQITPHCSQWIRTKLKAQORECASP 1500
Qy 1501 QHSABPIPTFKAKYILSDALLDKDSDSGVRSNESPNSHLHNEAADDSDOLEKANLI 1560
Db 1501 QHSABPIPTFKAKYILSDALLDKDSDSGVRSNESPNSHLHNEAADDSDOLEKANLI 1560
Qy 1561 ELEDEGHSGKRGMPHSLGLQDPIIARMSI CSEDKSPSECSLIASSPRESWPACOKAYN 1620
Db 1561 ELEDEGHSGKRGMPHSLGLQDPIIARMSI CSEDKSPSECSLIASSPRESWPACOKAYN 1620
Qy 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQNLKSMHAK 1680
Db 1621 LNRTPTSTVTLNNNTAPTNRANQNFDEIGIRETSQVILRPGSPNPTAVQNLKSMHAK 1680
Qy 1681 RSQRSSYTLRSLKDASELHHAASSESTGFGEERESIL 1715
Db 1681 RSQRSSYTLRSLKDASELHHAASSESTGFGEERESIL 1715

RESULT 2
Q9EQG6 PRELIMINARY; PRT; 1762 AA.
ID AC Q9EQG6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIDINS220.
GN Name=Kidins220;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568256; PubMed=10998417; DOI=10.1074/jbc.M005261200;
RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozengurt E.,
RA Schiavo G.;
RT Identification and cloning of Kidins220, a novel neuronal substrate
of protein kinase D.;
RL J. Biol. Chem. 275:40048-40056(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Mitchell M., Schiavo G.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF239045; AAC35185.2; -.
DR HSSP; P09959; 1SW6.
DR InterPro; IPR002110; ANK.
DR PFAM; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 10.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1762 AA; 195714 MW; 0CB2689A571F8AE4 CRC64;

Query Match 97.6%; Score 8639.5; DB 2; Length 1762;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1691; Conservative 6; Mismatches 15; Indels 53; Gaps 5;
Qy 1 MSVLISQSVINYVEENIPALKLEKCKDVDERNECGQTFPLMAAEQGNVEIVKELKN 60
Db 1 MSVLISQSVINYVEENIPALKLEKCKDVDERNECGQTFPLMAAEQGNVEIVKELKN 60
Qy 61 GANCNLEDLDNWTALLSASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLDNWTALLSASKEGHIHIVEELLKSGASLEHRDMGWTALMWACYKGRDVV 120

SQ	SEQUENCE	1777 AA; 197209 MW; B6505923FB45F143 CRC64;	
	Query Match	92.8%; Score 8212; DB 2; Length 1777;	
	Best Local Similarity	89.7%; Pred. No. 0;	
	Matches 1590; Conservative	53; Mismatches 70; Indels 60; Gaps 4;	
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAABQGNVVEIKELKN	60
Db	7	MSVLISQSVINYVEENIPALKALLEKCKVDVERNECGQTPMLAABQGNVVEIKELKN	66
Qy	61	GANCNLEDLNTWLTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMWACVKGRTDVV	120
Db	67	GANCNLEDLNTWLTALISASKEGHVHIVEELLKCGVNLHRDMGWTALMWACVKGRTDVV	126
Qy	121	ELLISHGANPSVTGLQVSVTPITWAAGRGHADIHILLQNGAKVNCSDKYGTTPLVWAAR	180
Db	127	ELLISHGANPSVTGL-QSVTPITWAAGRGHADIHILLQNGAKVNCSDKYGTTPLVWAAR	185
Qy	181	KHLECVKHLLAMGADVDOEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA	240
Db	186	KHLECVKHLLAMGADVDOEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA	245
Qy	241	LMIASKEGHIIEVQDLDDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300
Db	246	LMIASKEGHIIEVQDLDDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	305
Qy	301	QDNKTALYWAVERKGNATWDRDILQCNPDTEI CTYKGETPLIKATKORNIIEVVELLDKGA	360
Db	306	QDNKTALYWAVERKGNATWDRDILQCNPDTEICTYKGETPLIKATKORNIIEVVELLDKGA	365
Qy	361	KVSAVDKGGTPTLHVLRGRSRLAELLNPKDGRLLYRPNKAGETPNYIDCSHOKSIL	420
Db	366	KVSAVDKGGTPTLHVLRGRSRLAELLNPKDGRLLYRPNKAGETPNYIDCSHOKSIL	425
Qy	421	TOIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWSGSGSFLKKL	480
Db	426	TOIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWSGSGSFLKKL	485
Qy	481	EDEKTFAGOOTPLFOFSLHIVLFTLLCGGLVFAFPVDTNLAIAISFLALIYIF	540
Db	486	EDEKTFAGOOTPLFOFSLHIVLFTLLCGGLVFAFPVDTNLAIAISFLALIYIF	545
Qy	541	FIVYFGREGSNNWAWLSTRLARHIGVLELLPKLMFVNPPPELQOTTKALPVRFLL	600
Db	546	FIVYFGREGSNNWAWLSTRLARHIGVLELLPKLMFVNPPPELQOTTKALPVRFLL	605
Qy	601	TDYNRLSSVGGETSLSAEMIATLSDACEREFGLATLFRVFRTEESQKKKKWKTCCPLS	660
Db	606	TDYNRLSSVGGETSLSAEMIATLSDACEREFGLATLFRVFRTEESQKKKKWKTCCPLS	665
Qy	661	FVIFLFIIVGIIAGITILLAIIFRVDPKHLTWNAILISTASVVGAFVNLCTWQVLDLIL	720
Db	666	FVIFLFIIVGIIAGITILLAIIFRVDPKHLTWNAILISTASVVGAFVNLCTWQVLDLIL	725
Qy	721	NSQKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC	780
Db	726	NSQKRLHSAASKLHLKSGFMKVLKCEVELMARMAKTIDSFTQNTQRLVLIIDGLDAC	785
Qy	781	EQDKVLQMLDTRVLFVSKGPPIAIFASDPHIIKAINONLNSVLNDSNNGHDMRNIVH	840
Db	786	EQDKVLQMLDTRVLFVSKGPPIAIFASDPHIIKAINONLNSVLNDSNNGHDMRNIVH	845
Qy	841	LPVFLNSRGI SNARKFLVTSATNGDITCSDTTGTQEDTDRVQNSIGEMTKLGSKTALN	900
Db	846	LPVFLNSRGI SNARKFLVTSATNGDITCSDTTGTQEDTDRVQNSIGEMTKLGSKTALN	905
Qy	901	RRDYRRQMRQITTRQMSFDLTKLLVTEDFWSDISQPTWRRLLINIVSVTGRLLRANQIT	960
Db	906	RRDYRRQMRQITTRQMSFDLTKLLVTEDFWSDISQPTWRRLLINIVSVTGRLLRANQIS	965
Qy	961	FNWDRLASWINLTQWQPYRTSWLLIYLEETEGIPDQWTLKTMVERISKNIPTTKDVEPLL	1020
Db	966	FNWDRLASWINLTQWQPYRTSWLLIYLEETEGIPDQWTLKTMVERISKNIPTTKDVEPLL	1025

Qy	1021	EIDGDIRNEVELSRTPLVARDVKTFELPCTVNLPDKLREIIADYRAAREQINIGLAY	1080
Db	1026	EIDGDIRNEVELSRTPLVARDVKTFELPCTVNLPDKLREIIADYRAAREQISIGLAY	1085
Qy	1081	PPLPLEHEGPPRPSPGYSQSPASVCSSASFNPGPGVVSQPHSSYYSGLSGPQHPFYN--	1138
Db	1086	PPLPLEHEGPPRPSPGYSQSPASVCSSASFNPGPGVVSQPHSSYYSGMTGPOHPFYNRP	1145
Qy	1139	-----RAAVP	1143
Db	1146	FFAPYLYTPRYPPGSGHILSRPSVKTSLPRDQNGGLEVIKEDAEGLSPSTDSRSGSP	1205
Qy	1144	ATGSSLLLSMTVDVCEKLRQIEGLDQNMPOYCTTIKKANINGVLSOCNIDELKKEM	1203
Db	1206	APGFVLLNSLNDVACEKLRQIEGLDQNMPOYCTTIKKANINGVLSOCNIDELKKEM	1265
Qy	1204	AMNFGDWHLFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELS	1263
Db	1266	NMFGDWHLFRSIVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHTELS	1325
Qy	1264	SQPYTLNFSFELNLTGLDEGAPRHSNLSWQSQTRTPSLSLNSQDSIESIKLTDKV	1323
Db	1326	SQPYTLNFSFELNLTGLDEGAPRHSNLSWQSQTRTPSLSLNSQDSIESIKLTDKV	1385
Qy	1324	QAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTYVYGSSSGGSIHSTLEQERKEG	1383
Db	1386	QAEYRDAYREYIAQMSOLEGGTSGSTISGRSSPHSTYVYGSSSGGSIHSTLEQERKEG	1445
Qy	1384	ELKQEDGRKSFMLKMGDVIDYSSGVSSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER	1443
Db	1446	EPKPDGRKSFMLKMGDVIDYSSGVSSTNEASPLDPIITEDEKSDQSGSKLLPGKKSER	1505
Qy	1444	PSLQFQDQLKKGGLRYQKLPDEDESGTGRVOITPHCSKMIRTKLKAKQRCACAPQEH	1503
Db	1506	SSLFQDQLKKGGLRYQKLPDEDESGTGRVOITPHCSKMIRTKLKAKQRCACAPQEH	1564
Qy	1504	SAEPIRTFKAKBYLSDALLDKDSDSGVRSNESPNSHNEAADDQLEKANIIELE	1563
Db	1565	SAEPIRTFKAKBYLSDALLDKDSDSGVRSNESPNSHNEAADDQLEKANIIELE	1624
Qy	1564	DEHSGKRGMPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPEESWPACQAYNLNR	1623
Db	1625	DDSHSGKRGMPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPEESWPACQAYNLNR	1684
Qy	1624	TPSTVTLNNTAPTNRANONFDEIGRETSQVILRPGSPNPNTAVONENLKSVAHKRSQ	1683
Db	1685	TPSTVTLNNTAPTNRANONFDEIGRETSQVILRPGSPNPNTAVONENLKSVAHKRSQ	1744
Qy	1684	RSSYTRLSKDASELH-AASSESTGFGEEESIL	1715
Db	1745	RSSYTRLSKDASELH-AASSESTGFGEEESIL	1777

RESULT 4

Q80TG7	PRELIMINARY;	PRT; 1693 AA.
ID	Q80TG7	Q80TG7;
AC	Q80TG7	01-JUN-2003 (T-EMBLrel. 24, Created)
DT	01-JUN-2003	(T-EMBLrel. 24, Last sequence update)
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)
DE	MKIAA1250	protein (Fragment).
GN	Name=C330002119Rik; Synonyms=mkIAA1250;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Brain;	
RX	MEDLINE=22579291; PubMed=12693553;	
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,	
RA	Nakajima D., Nagase T., Ohara O., Koga H.;	

RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 DR EMBL; AK122478; BAC65760.1; -.
 DR HSSP; P09959; 18W6.
 DR MGD; MGI:1924730; C330002119Rik.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 10.
 DR PRINTS; PR01415; ANKTRIN.
 DR SMART; SM00248; ANK; 9.
 DR PROSITE; PS50088; ANK_REPEAT; 8.
 DR PROSITE; PS50297; ANK_REPEAT; 1.
 KW ANK repeat.
 FT NON TER
 SQ SEQUENCE 1 1693 AA; 188253 MW; 599DE2A3419D4C9D CRC64;

Query Match 89.1%; Score 7892; DB 2; Length 1693;
 Best Local Similarity 90.1%; Pred. No. 0;
 Matches 1544; Conservative 30; Mismatches 39; Indels 100; Gaps 6;

QY 83 HIHIVEELLKSGASLEHRDMGCGTALMWACYKGRDVTVELLSHGANGPSVTGLQYSVYPI 142
 DB 1 HIHIVEELLKCGANLEHRDMGCGTALMWACYKGRDVTVELLSHGANGPSVTGL-QYSVYPI 59
 QY 143 IWAAGRHADIVHLLQNGAKVNSDKYGTTPVWAARKHLECVKHLANGADVQEGA 202
 DB 60 IWAAGRHADIVHLLQNGAKVNSDKYGTTPVWAARKHLECVKHLANGADVQEGA 119
 QY 203 NSMTALIVAVKGGYTQSVKEIKENPNVNLTKDQNTALMIASKEGHEIVQDILLDAGTY 262
 DB 120 NSMTALIVAVKGGYTQSVKEIKENPNVNLTKDQNTALMIASKEGHEIVQDILLDAGTY 179
 QY 263 VNIIPDRGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATVYRDI 322
 DB 180 VNIIPDRGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTALYWAVEKGNATVYRDI 239
 QY 323 LQCPNPTETICQGETPLIKATKRNIEVVELLDKAKYSAVDKKGDTPLHVAIRGRSR 382
 DB 240 LQCPNPTETICQGETPLIKATKRNIEVVELLDKAKYSAVDKKGDTPLHVAIRGRSR 299
 QY 383 RLAEALLRNPKDGLLYRPNKAGETPNIDCSHQKSLITQIFGARHLSPTETDGMGLGYD 442
 DB 300 RLAEALLRNPKDGLLYRPNKAGETPNIDCSHQKSLITQIFGARHLSPTETDGMGLGYD 359
 QY 443 LYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLEDMKTFAGQQTPLFPQSWLI 502
 DB 360 LYSSALADILSEPTMQPPICVGLYAQWGSGKSFLLKKLEDMKTFAGQQTPLFPQSWLI 419
 QY 503 VFLTLCCGGLGVFAFVDTNLAIAISLSFLALYIFFVIYFGGRREGESNNWAWALS 562
 DB 420 VFLTLCCGGLGVFAFVDTNLAIAISLSFLALYIFFVIYFGGRREGESNNWAWALS 479
 QY 563 TRLARHIGYLELLEFKLMFVNPPPEQTTKALPVRFLETDVNRSLSSVGGTSLAEMIATL 622
 DB 480 TRLARHIGYLELLEFKLMFVNPPPEQTTKALPVRFLETDVNRSLSSVGGTSLAEMI --- 536
 QY 623 SDACERFGLFATLRFVFRTEESQGGKKWKTCCLPSFVIFLFIIVGCIITLAIIFR 682
 DB 537 -----FEVLRLLKILRV-----KXRGKTCCLPSFVIFLFIIVGCIITLAIIFR 581
 QY 583 VDPKHLTNAILISIASVGLAFVLCRTWQVNLDSLLNSQKRLHSAASKHLKXSEGF 742
 DB 582 VDPKHLTNAILISIASVGLAFVLCRTWQVNLDSLLNSQKRLHSAASKHLKXSEGF 641
 QY 743 MKVLKCEVELMARWAKTIIDSTQNTQRLVLIIDGLDACEQDKVLQMLDTRVLFSGKPGFI 802
 DB 642 MKVLKCEVELMARWAKTIIDSTQNTQRLVLIIDGLDACEQDKVLQMLDTRVLFSGKPGFI 701
 QY 803 AIFASDPHIIITKAINQNLSVLRDSNINGHDYMRNIVHLPVFLNSRGLSNARKFLVTSAT 862
 DB 702 AIFASDPHIIITKAINQNLSVLRDSNINGHDYMRNIVHLPVFLNSRGLSNARKFLVTSAT 761
 QY 863 NGDITCSDTTCTQEDTDRRVSONSLGEMTKLGSXTALNRRDTYRRROMORTITROMSFDL 922
 DB 762 NGDISCSEATGVEDADRVSQSLGEMTKLGSXTALNRRDTYRRROMORTITROMSFDL 821
 QY 923 TKLLVTEDFWSDISFQPMRRLNIVSVTGRLLRANQITFNWDRLASWLNLTQEWPYRTSW 982
 DB 822 TKLLVTEDFWSDISFQPMRRLNIVSVTGRLLRANQITFNWDRLASWLNLTQEWPYRTSW 881
 QY 983 LILYLETEGLPDQMTLTKTYERISKNIPTTKQVEPLEIIEIDGDIRNFVFSSRTPVLVA 1042
 DB 882 LILYLETEGLPDQMTLTKTYERISKNIPTTKQVEPLEIIEIDGDIRNFVFSSRTPVLVA 941
 QY 1043 RDVKTFLPCTVNLDPKLEIITADVRAAREQINIGLAYPPLPLHEGPPRPSGYSQASV 1102
 DB 942 RDVKTFLPCTVNLDPKLEIITADVRAAREQINIGLAYPPLPLHEGPPRPSGYSQASV 1001
 QY 1103 C-SSASFNPGPPGGVSPQPHSSYISGLSGPQHPFYNR----- 1139
 DB 1002 CSSASFNPGPPGGVSPQPHSSYISGLSGPQHPFYNRPFAPYLYTPRYYPGSGSHLIS 1061
 QY 1140 -----AAV 1142
 DB 1062 RSSVKTSLPRDQNNGLSEVIKEDAAEGLPSPPTASSREKSWTRKQLMELCDSGFNKQASV 1121
 QY 1143 PATGSSLLLSMTVDVVCCKLROIIEGLDONMMPOYCTTIKKANINGRVLSCNIDELKKE 1202
 DB 1122 PATGSSLLLSMTVDVVCCKLROIIEGLDONMMPOYCTTIKKANINGRVLSCNIDELKKE 1181
 QY 1203 MAMNFGDHLFRSNVLEMRSVESQVPEPRFLNENSSAPVPHGESARRSHTLPTLTEL 1262
 DB 1182 MAMNFGDHLFRSNVLEMRSVENQVPEPRFLNENSSAPVAHGESARRTSSELPTLTEL 1241
 QY 1263 SSQTPYTLNFSFEBELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSINSQDSSEIRISKLTDK 1322
 DB 1242 SSQTPYTLNFSFEBELNTLGLDEGAPRHSNLSWQSQTRRTPSLSSINSQDSSEIRISKLTDK 1301
 QY 1323 VQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYIYIGQSSSGSIHSLTQERKKE 1382
 DB 1302 VQAEYRDAYREYIAQMSQLEGGTGSSTISGRSSPHSTYIYIGQSSSGSIHSLTQERKKE 1361
 QY 1383 GELKQEDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPTTBEDEKSDQSGSKLLPGKKSSE 1442
 DB 1362 SELKQEDGRKSFMLKRGDVIDYSSSGVSTNEASPLDPTTBEDEKSDQSGSKLLPGKKSSE 1421
 QY 1443 RPSLFQTDLKGGLRYQKLPDEDESGTGRVOITPHCSNMIRTKLAKAKORCASQOE 1502
 DB 1422 RPSLFQTDLKGGLRYQKLPDEDESGTGRVOITPHCSNMIRTKLAKAKORCASQOE 1480
 QY 1503 HSAEPIRTFKAEYLSDALDKDSDGVRNNESSPNHSLHNEAADDSDLEKANLTEL 1562
 DB 1481 HSAEPIRTFKAEYLSDALDKDSDGVRNNESSPNHSLHNEAADDSDLEKANLTEL 1540
 QY 1563 EDEGHSGKRWPHSLSGLOPIIARMSICSEDKKSPSECSLIASSPESWPAQKAYNLN 1622
 DB 1541 EDEGHSGKRWPHSLSGLOPIIARMSICSEDKKSPSECSLIASSPESWPAQKAYNLN 1600
 QY 1623 RTPSTVTNNNTAPTNRANQNFDEIGRETQSQVILRPGSPNPTAVQENLKSMAHKRS 1682
 DB 1601 RTPSTVTNNNTAPTNRANQNFDEIGRETQSQVILRPGSPNPTAVQENLKSMAHKRS 1660
 QY 1683 QRSSTYLSKASSELHAASSESTGFGBERESIL 1715
 DB 1661 QRSSTYLSKASSELHAASSESTGFGBERESIL 1693

RESULT 5
 Q80SX9
 ID Q80SX9 PRELIMINARY; PRT; 1554 AA.
 AC Q80SX9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)


```
Db 1261 TEEDEKSDQSGSKLLPGKKSSERSPSLFTQDULKXGSLRYOKLPSDEDESGTESDNTPT- 1319
Qy 1481 CSKMRTRKLRKAKQREKASPOEHSABEPIRTPIKAKYLSLALLDKOSSDGVSRNNESSP 1540
Db 1320 LLKDDKOKKAGSKAERVAKSPHSVEPIRTPIKAKYLSLALLDKOSSDGVSRNNESSP 1379
Qy 1541 NLSLHNEAADSDQLEKANLIELEDEGHSGKRGMPHSLSGLQDDPIIARMSICSEDKKSPSE 1600
Db 1380 NLSLHNEAADSDQLEKANLIELEDEGHSGKRGMPHSLSGLQDDPIIARMSICSEDKKSPSE 1439
Qy 1601 CSLIASSPEESWPACOKAYNINLRPTSTVTLNNNTAPTNRANQNDELEGRETQSIVLRP 1660
Db 1440 CSLIASSPEESWPCQKAYNINLRPTSTVTLNNNTAPTNRANQNDELEGRETQSIVLRP 1499
Qy 1661 GPSNPNTAVQENLKSMAHKRSQSSYTRLKSDASELHAASESTGFGEERESIL 1715
Db 1500 GPSNPNTAVQENLKSMAHKRSQSSYTRLKSDASELHAASESDTGFGEERESIL 1554

RESULT 6
Q7T163
ID Q7T163 PRELIMINARY; PRT; 1680 AA.
AC Q7T163;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SI:d2119J18.2 (Novel protein similar to rat kinase D-interacting
DE substate of 220 kDa (KIDINS220) ).
CN Name:SI:d2119J18.2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL954721; CAELI7588.1; -.
DR HSSP; P42773; 1IHB.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00023; Ank; 11.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_1.
KW ANK repeat; Kinase.
SQ SEQUENCE 1680 AA; 185948 MW; 4F8AB3377DD5E60 CRC64;

Query Match 70.4%; Score 6233.5; DB 2; Length 1680;
Best Local Similarity 71.8%; Pred. No. 7.3e-315;
Matches 1243; Conservative 180; Mismatches 232; Indels 77; Gaps 26;

Qy 1 MSVLISQVINYVEENIPALKALKERKQVDERNECGQTPLMLAABQGNVZIVKELLKN 60
Db 9 MTTLAIQNLFSYVEENLAAVKVHLDFKFEVDGRSDNGQTFPLMLASQGSLEIVQELIRR 68
Qy 61 GANCNLEDLNWTALISASKEGHHIYVELLKSASLEHHRDMGWTALMWACYKGRDVTY 120
Db 69 GANYNLDDVDCWSALISAKEGHEVYVVKELLSENSAYIEHRDMGWTALTWASYKGRVEVA 128
Qy 121 ELLLSHGANSVTCLOYSVYPIIWAAGRGHADIYHLLLLONGAKVNCSDKYGTTPLVWAAR 180
Db 129 TVLLENGANPNNTQQQSVYPIIWAAGRGHAEIVKLLLEHGAKVNCSDKYGTTPLIWAAR 188
Qy 181 KGHLECVKHLIAGADVQDEGANSMTALIVAVKGGYTSQVKELIKRPNVNLTKDGNTA 240
Db 189 KGHYDCVWHLENGADVQDEGANSMTALIVAVKGGYTEVVKELIKRPNVNLTKDGNTA 248
Qy 241 LMIASKEGHIEIVQDLIDAGTYVNIIPDRSGDTVLIGAVRGGHVIEVALLQKYADIDIRG 300
Db 1261 TEEDEKSDQSGSKLLPGKKSSERSPSLFTQDULKXGSLRYOKLPSDEDESGTESDNTPT- 1319
Qy 1481 CSKMRTRKLRKAKQREKASPOEHSABEPIRTPIKAKYLSLALLDKOSSDGVSRNNESSP 1540
Db 1320 LLKDDKOKKAGSKAERVAKSPHSVEPIRTPIKAKYLSLALLDKOSSDGVSRNNESSP 1379
Qy 1541 NLSLHNEAADSDQLEKANLIELEDEGHSGKRGMPHSLSGLQDDPIIARMSICSEDKKSPSE 1600
Db 1380 NLSLHNEAADSDQLEKANLIELEDEGHSGKRGMPHSLSGLQDDPIIARMSICSEDKKSPSE 1439
Qy 1601 CSLIASSPEESWPACOKAYNINLRPTSTVTLNNNTAPTNRANQNDELEGRETQSIVLRP 1660
Db 1440 CSLIASSPEESWPCQKAYNINLRPTSTVTLNNNTAPTNRANQNDELEGRETQSIVLRP 1499
Qy 1661 GPSNPNTAVQENLKSMAHKRSQSSYTRLKSDASELHAASESTGFGEERESIL 1715
Db 1500 GPSNPNTAVQENLKSMAHKRSQSSYTRLKSDASELHAASESDTGFGEERESIL 1554
```

```
Db 249 LMIAAKEGYTEIVQDLIDAGTYVNIIPDRSGDTVLIGAVRGGHVIEVALLHXYADIDIRG 308
Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLLDKGA 360
Db 309 QENKTALYWAVEKGNATWVRDILQCNPDTEITTKDSETPLIKATKMRNIEVVELLLDKGA 368
Qy 361 KVSADVKKGTPTLHVATRGSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQSKIL 420
Db 369 KVSADVKKGTPTLHVATRGSRRLAELLLRNPKDGRLLYRPNKAGETPYNIDCSHQSKIL 428
Qy 421 TQIFGARHLSPTEDEGMGLGYDLYSSALADILSPTEQPPICVGLYQAQWGSGKFLKKL 480
Db 429 TQIFGARHLSPTEDEGMGLGYDLYSSALADILSPTEQPPICVGLYQAQWGSGKFLKKL 488
Qy 481 EDEKMTFAGQTEPLFQFWSMLIVFLTLCLCGGLGVFAFPVDNTLAIATISLFLAIYIF 540
Db 489 EDEKMTFAGQTEPLFQFWSMLIVFLTLCLCGVALVIGFTVDPKLAATISLIALLYVF 548
Qy 541 FVIYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRPLF 600
Db 549 FVVYFGRRREGESWNWALSTRLARHIGYLELLFKLMFVNPPELPEQTTKALPVRPLF 608
Qy 601 TDYNLSSVGETSLAEIATLSDACEREGFLATRLFRVPRTEESQKKKKKTKCLPS 660
Db 609 TDYNLSSVGETSMAEMIATLSDACEREGFLATRLFRVFKTEDTQKKKKKTKCCIPS 668
Qy 661 FVIFLFTVGCIIAGITLILAFRVDPKHLTVNAILISIASVVLAFVLCNRTWQVOLSLL 720
Db 669 FVIFLFTVGCIIAGITLILAFRVDPKHLTVNAILISIASVVLAFVLCNRTWQVOLSLL 728
Qy 721 NSQKRLHSAASKLHLKLGSEGMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDAC 780
Db 729 NSQKRLHSAANKHKLKLGSEGMKVLKCEVELMARMARKTIDSFQNTQRLVVIIDGLDAC 788
Qy 781 EODKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINQNLNSVLNRDNINGHDYMRNIVH 840
Db 789 EODKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINQNLNSVLNRDNINGHDYMRNIVH 848
Qy 841 LPVFLNSRGLSNARKPLVTSATNGDICTSDTTGTQEDTDRRVSONSLGEMTKLASKTALN 900
Db 849 LPVFLNSRGLSSAKKMCAPAPANGE--TGNSEGHHELDRLKLSQNSLODQTKFSKTYLN 906
Qy 901 RRDYRRRQRTITRQMSFDLTKLVTEDFWDFDISPQTMRRLLNIVSVTGRLLRANQIT 960
Db 907 RRDYRRRQMSVTRQMSFDLTKLVTEDFWDFDISPQTMRRLLNIVSVTGRLLRANQIS 966
Qy 961 FNVDRLASWINLTEOMPYRTSWLLIYLEETEGLPDQMTLKTMYERIKNPTTKDVEPLL 1020
Db 967 FNVDRLASWINLTEOMPYRTSWLLIYLEETEGLPDQMTLKTMYERIKNPTTKDVEPLL 1026
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGLAY 1080
Db 1027 EIDGDIRNFVFLSSRTPVLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGLAY 1086
Qy 1081 PPLPLHEGPPRPSPGYSQSPASVCS-SASFNPGPPPGVSPQPHSSYSYSGLSGPPHFPYNR 1139
Db 1087 PTLPLQSG--RPISMYSQSSACSPTASFNGPYNPPCGVSPQPHSAYFSGMAGPQHPFYNR 1144
Qy 1140 --RAVPATGCSLLISLMSVTVDVCEKLQIIEGLDONMMPQCTTIKKANINGRVLSCQDID 1197
Db 1145 GSASVSGTFSILLSSMSTDVI CERVLIIDGIDQNLISQYTATIKKANINGRVLSCQDID 1204
Qy 1198 ELKKEMANFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSHTL 1257
Db 1205 ELKKEMANFGDWHLPFRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGSARSSHTL 1262
Qy 1258 PLTESSQTP- YTLNFSFEELNTLGLDEGAPRHSNLSWQSQTRTPSLSSINSQSSSIEI 1316
Db 1263 GVAGNTDTSMPYNFLSFEELSNGVLEE-PPRHVNATMGTTHRTPSWSSINSQSSNEI 1321
Qy 1317 SKLTKDQVARYDAYREYIAQMSOLEGGTGSSTISGRSSPHSTYIYICQSSSGGSIHSTLE 1376
Db 1322 CKLTKQQAERYNAYEDIYASMSQLE-----LGMKEKPVPP-FVSQ-----LMHSSSE 1367
```

QY 1377 QERKGEKQEDGRKSFMLKRG-----DVIDYSSGVSTNEASPLDPTITEDEKSDQS 1430
Db 1368 DKK-KDG--NDQGRKS-VSKRGSTSGSDNTDYASA-----DAATLDPTITEEDKVDHG 1418
QY 1431 GSKLLPGKSS-BRPSLFQ-TDLKLG-GGLRYQKLPSPDESGTGRVQITTPHCSRWIRT 1487
Db 1419 SSKSLGRKTSKGVSLFQADLKLKAGGSRQKLTSDDEES-----EESDNAPLLKDG 1473
QY 1488 KRLKAKORECASPQERSAEPITPIKAVEYLSALLDKKSDSGSVRSNESSPNHSHNE 1547
Db 1474 KPEAKASDGD-----RSLTGKQVLS-----DKDSSDSGVSRSNESSPNHSLQDE 1520
QY 1548 AADSQLEKANLIELEDEGSGRGWPHSLSGLODPTIARMSTCSDEKSPSPCSLIASS 1607
Db 1521 EADLSOSERANLIEDEESARKRGLPNSLUGLODPTIARMSTCSDE-----QCSLIASS 1575
QY 1608 PEESWAPQKAYNLNRTPTSTVTLNNNT--APTNRANQNFDEIGIRET--SOVILRPGPS 1663
Db 1576 PEESWPS-SKSYNLNRTPTSTVTLNNNTAQGNHIRQPSDSSNTTSTTTCSDVIINPTGS 1634
QY 1664 PNTAVQNEHLKSMHAKRSORSYTRLSKDASELHAASSESTFGBERSIL 1715
Db 1635 TTSATQTNENVRVHLKRLNPG-----DPPILKVSSETVTFGBERSIL 1690

RESULT 7
Q6P7Y1
ID Q6P7Y1 PRELIMINARY; PRF; 1672 AA.
AC Q6P7Y1;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypochemical protein zgc:63531.
GN ORFNames=zgc:63531;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Srausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061450; AH61450.1;
DR HSSP; Q60773; 1AP7.
DR ZFIN; ZDB-GENE-030131-7824; zgc:63531.
DR InterPro; IPR002110; ANK.

DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00023; Ank; 10.
DR PRINTS; PRO1445; ANKTRIN.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS0088; ANK REPEAT; 10.
DR PROSITE; PS0297; ANK REP REGION; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 1672 AA; 185035 MW; 75BCF10280B410B CRC64;

Query Match 70.3%; Score 6219.5; DB 2; length 1672;
Best Local Similarity 71.7%; Pred. No. 3.8e-314;
Matches 1241; Conservative 180; Mismatches 234; Indels 77; Gaps 26;

QY 1 MSVLISQSVINVEENIPALKALLEKCKVDVERNECGQTPLMLAAEQGNVEIKVLLKN 60
Db 1 MTTATLQNLFSYVEEENLAHVLDKPKVEVDGRSDNGQTPLMLASEQSLVQSLIRR 60
QY 61 GANCNLEDDNNTALISASKEGHIHIVEELLKSGASLEHRDMGWTALMAWACYKGRDVTV 120
Db 61 GANVNLDVDDCSALISAAKEGHEVVEVKELSENSAYIEHRDMGWTALTWASYKGRVEVA 120
QY 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADIHLLQNGAKVNCSDKYGTTPLYWAAR 180
Db 121 TVLLENGANPNTTGGQYSVYPIIWAAGRGHABIVKLLLEHGAQVNCSDKYGTTPLYWAAR 180
QY 181 KGHLECVKLLAMGADVDOEGANSMTALIVAVKGGYTSQSVKELLKNPNVNLTKDGNTA 240
Db 181 KGHYDCWMLLENGADVDOEGANSMTALIVAVKGGYTSQSVKELLKNPNVNLTKDGNTA 240
QY 241 LMAASKEGHEIIVQDLIDAGTVNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDRG 300
Db 241 LMAAKEGYEIVQDLIDAGTVNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDRG 300
QY 301 QNKATLYWAVEKGNATWVRDILQCNPDTEICTKQGETELIKATKRNTEVVELLDKGA 360
Db 301 QENKATLYWAVEKGNATWVRDILQCNPDTEITTKQSETELIKATKRSIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLVRPNKAGETPNVDCSHQSKIL 420
Db 361 KVSADVKKGDTPLHVAIRGSRRLAELLRNPKDGLLVRPNKAGETPNVDCSHQSKIL 420
QY 421 TQIFGARHLSPETDGDMLGYDLYSSALADILSEPTMQPPICVGLYAQSGSKSFLKKKL 480
Db 421 TQIFGARHLSPESDGDMLGYDLYSSALADILSEPTMQPPICVGLNTQSGSKSFLKKKL 480
QY 481 EDEMTFAQQQTEPLFQFSWLVFVTLCCGGLGVFAPVDVTNLAIALISLFLALIYF 540
Db 481 EDEMTFAQQQTEPLFQFSWLVFVTLCCGGLGVFAPVDVTNLAIALISLFLALIYF 540
QY 541 FVIVYFGGRREGESNNWALSTRLARHIGYLELLEFKLMFVNPPPEQOTTALPVRELF 600
Db 541 FVIVYFGGRREGESNNWALSTRLARHIGYLELLEFKLMFVNPPPEQOTTALPVRELF 600
QY 601 TYNRLSSVGGTSLAEMIATISDACEREFGLATRLFRVFRTERESQGGKKWKTCCLPS 660
Db 601 TYNRLSSVGGTSLAEMIATISDACEREFGLATRLFRVFRTERESQGGKKWKTCCLPS 660
QY 661 FVIFLPIVCGITAGITLLAIFRVPDKHFTVNAILISIASVWGLAFVLCRTWQVLDL 720
Db 661 FVIFLPIVCGITAGITLLAIFRVPDKHFTVNAILISIASVWGLAFVLCRTWQVLDL 720
QY 721 NSQKRLHSAASKLHKSEGEKMKVKCEVELMARMAKTIDSTFQNTOTRLVILIDGLDAC 780
Db 721 NSQKRLHSAASKLHKSEGEKMKVKCEVELMARMAKTIDSTFQNTOTRLVILIDGLDAC 780
QY 781 EQDKVLQMLDVTVRVLFSGKPFIAIFASDPHIIKAINQNLNSVLRDSNNGHYDMEIVH 840
Db 781 EQDKVLQMLDVTVRVLFSGKPFIAIFASDPHIIKAINQNLNSVLRDSNNGHYDMEIVH 840
QY 841 LPVFLNSRGLSNARPLVTSATNGDITCSDTGTGTQEDTDRRVDSQNSLGEMTKLGSALTAN 900
Db 841 LPVFLNSRGLSNARPLVTSATNGDITCSDTGTGTQEDTDRRVDSQNSLGEMTKLGSALTAN 900

Qy	901	RDYTRRRQMTTITRQMSFDLTLLIYVTEWFSDISPQWTRRLNIIVSVTGRLLRANQIT	960		
Db	999	RDYTRRRQMTTITRQMSFDLTLLIYVTEWFSDISPQWTRRLNIIVSVTGRLLRANQIS	958		
Qy	961	FNWDRLASWNLTEQWPYRYSWLLIYLEEETGLPDQMTLKTMYERISKNIPTTKDVEPLL	1020		
Db	959	FNWDRLASWNLTEQWPYRYSWLLIYLEEETGLPDQNLKTIYERISKNIPTTKDVEPLL	1018		
Qy	1021	EIDGDIRNFVFLSSRTPLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGGLAY	1080		
Db	1019	EIDGDIRNFVFLSSRTPLVARDVKTFLPCTVNLDPKLREIIADVRAAREQINIGGLAY	1078		
Qy	1081	PPLPLHGGPPRPPGGYQOPASVCS-SASFNGPPFGVVPSPQPHSSYSGLSGQHPFPYNR	1139		
Db	1079	PTLPLOEG--RPISMYQSQSACPTASFGFYPNPGVSPQPHSAFSGMAGQHPFPYNR	1136		
Qy	1140	--AAPTATGSSLSSMTVDVVCCKLRQIIEGLDONMMPQVCTTIKKANINGRVLSCNID	1197		
Db	1137	GSASVSGVTSILLSSMTDVCICRVKLIIDGIDONLISQVYATIKKANINGRVLSCNID	1196		
Qy	1198	ELKEMAMNFGDMHLFRFSMWLEKRSVBSQVVPDPREFLNENSSAPVPHGESARRSHTEL	1257		
Db	1197	ELKEMAMNFGDMQLFRTSVLEHVENQVLEHAP--SEQGSITVGHVPCFHAGAAQ	1254		
Qy	1258	PLTESSQTP-YTNFSGFEELNTLGLDEGAPRHSNLISQSQTRTTPSLSSLSNQSISIEI	1316		
Db	1255	GVAGNTDTPSPYNFNLSEELSNVGLSE-PPRHVNATWGTATHTPMSLSNQSSENEI	1313		
Qy	1317	SKLTDKVOAERDAYRYIAQMSOLEGTCGSSITSGRSSPHSYIYIGQSSSGSIHSTLE	1376		
Db	1314	CKLTDKQOAEYRNAYEDIYIASMSQLE-----LGMKEKVPVP-FVSQ-----LMHSSE	1359		
Qy	1377	QERKEGELKQEDGRKSFLEMKRG-----DIVDYSVGSTNEASPLDPTTEDEKSDQS	1430		
Db	1360	DKK-KDG--NDQGRKS-VSKRGSXSGSNTDYASA-----DAATLDPITEDEKVDHG	1410		
Qy	1431	GSKLLPGKKS-ERPSLFQ-TDLKLK-GGGLRYQKLPFSDDESGTGRVQITPHCRSMRT	1487		
Db	1411	SSKSLGRKTS GDKVSLFQGAADLKLAGGSGRYKLTSDDEES-----BEDNAPLLKDG	1465		
Qy	1488	KRLKAKQREKASQEHSAEPIRFIAKAYLSDALLDKXSSDSGVRNNESSPNHSIHNE	1547		
Db	1466	KXPEKASDGD-----RSLTKGKLYS---DKKSDSDGVRNNESSPNHSIQDE	1512		
Qy	1548	AADDSQLEKANLIELEDEHSGRKGMPHSISGLQDPTIARWSTICSDKKSPFCSLIASS	1607		
Db	1513	EADLSQSERANLIELEDENSARRKGLPNSLSGLQDPTIARWSTICSD-----QCSLIASS	1567		
Qy	1608	PEESWPACQKAYNLNRTPPSTVTLNNNT--APTNRANONFEIEGIRET--SQVILRPGPS	1663		
Db	1568	PEESWPS-SKSYNLNRTPPSTVTLNNNTAQQNHIRQPSDSNNTTSTTTCSDVINPGTS	1626		
Qy	1664	PNPTAVONENLKSMAHKRSORSSYTRLSKDASELHAASSSTGFGBERESIL	1715		
Db	1627	TTSATTQENVRVVHLKRLNPG-----DPEILKVSSETVTVFGBERESIL	1672		
RESULT 8					
ID	Q9UF42	PRELIMINARY;	PRT; 1031 AA.		
AC	Q9UF42;				
DT	01-MAY-2000	(T=EMBLrel. 13, Created)			
DT	01-MAY-2000	(T=EMBLrel. 13, Last sequence update)			
DT	01-MAR-2004	(T=EMBLrel. 26, Last annotation update)			
DE	Hypothetical protein DKFp2434F0621.				
GN	Name=DKFp2434F0621.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	[1]				
RN	SEQUENCE FROM N.A.				

[illegible]

```
Db 781 EQDKVLQMLDVTVRVLSKGFPIAFASDPHIIKAINQNLNSVLRDSNINGHDYMRNVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVQNSIGLGMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDSTTGIOEDADRVSQNSIGLGMTKLGSKTALN 900
QY 901 RRDYRRRQWRITTRQMSFDLTKLVTEWDFSDISPTQMRRLNLTNVSVTGRLLRANQIT 960
Db 901 RRDYRRRQWRITTRQMSFDLTKLVTEWDFSDISPTQMRRLNLTNVSVTGRLLRANQIS 960
QY 961 FNDRLASWINLAEQMPYRTSWLILYLETEGLPDQWTLKTYER-----IS 1007
Db 961 FNDRLASWINLAEQMPYRTSWLILYLETEGLPDQWTLKTYERCCGADSCDRDRIGIS 1020
QY 1008 KNI 1010
Db 1021 KSV 1023

RESULT 9
Q72322 PRELIMINARY; PRT; 691 AA.
AC Q72322;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFp686H14204 (Fragment).
GN Name=DKFp686H14204;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human fetal brain;
RA Lauber J.; Bahr A.; Mewes H.W.; Weil B.; Amid C.; Osanger A.; Fobo G.;
RA Han M.; Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX38189; CAD98059.1; -
DR InterPro; IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON_TER
FT NON_TER
SQ SEQUENCE 691 AA; 75760 MW; F5783F24DDEA6027 CRC64;

Query Match 35.08; Score 3097; DB 2; Length 691;
Best Local Similarity 86.8%; Pred. No. 1.8e-152;
Matches 601; Conservative 32; Mismatches 57; Indels 2; Gaps 2;

QY 1025 DIRFEVFLSSRTPVLVARDVKTFLPCTVNLDPKRLREIADVRAAREQINIGGLAYPPLP 1084
Db 1 DIRFEVFLSSRTPVLVARDVKVFLPCTVNLDPKRLREIADVRAAREQISIGGLAYPPLP 60
QY 1085 LHGGPPRPPGYSOPASVCSSAFNGFPFGVSPQPHSSYSGLSGPPQHPFNRAAVPA 1144
Db 61 LHGGPPRPPGYSOPSPVCSSTSFNGFPAGVSPQPHSSYSGMTGTPQHPFNRRSGGPA 120
QY 1145 TGSLLSSMTVDVVCCKLQIQLDONMMPOYCTTIKKNINGRVLSCQNIDELKKEWA 1204
Db 121 PGVPVLNLSNDAVCEKQIQLDONMMPOYCTTIKKNINGRVLSCQNIDELKKEWA 180
QY 1205 MNFGDMLFRSMVLEKRSVSVQVPDPFLNENSSAPVPHGSARRSSHTLPLTELSS 1264
Db 181 MNFGDMLFRSTVLEMAESHVVPDPFLSSSSGPA PHGPARRAASHNELPHTLSS 240
QY 1265 QTPYTLNFSPEELNTLGLDGAPRHSLSWQSOTRRTPSLSSNQSSSIEISKLTDKVQ 1324
Db 241 QTPYTLNFSPEELNTLGLDGAPRHSLSWQSOTRRTPSLSSNQSSSIEISKLTDKVQ 300
QY 1325 AEYRDAYREYIAQMSOLEGGTGSSTISGRSSPHSTYVIGOSSGGSHTSLEBERKEGE 1384
Db 301 AEYRDAYREYIAQMSOLEGGTGSSTISGRSSPHSTYVIGOSSGGSHTSLEBERKEGKDS 360
QY 1385 LKQEDGRKSLMKRGDVIDYSSSGVSTNEASPLDPITEDEKSDQSGSKLLPGCKSSERP 1444
```

```
Db 361 PKPDDGRKSLMKRGDVIDYSSSGVSTNEASPLDPITEDEKSDQSGSKLLPGCKSSERS 420
QY 1445 SLFQTDLKLKGGGLRYQKLPSPDESDSGTRVQITTHCSKMTKRLKAKORECASQEHHS 1504
Db 421 SLFQTDLKLKGGGLRYQKLPSPDESDSGTRVQITTHCSKMTKRLKAKORECASQEHHS 479
QY 1505 ASPIRTFIKAKVLSDALLDKKDSSDGVRSNESPNSHNEAADDSOLEKANLIELED 1564
Db 480 ASPIRTFIKAKVLSDALLDKKDSSDGVRSNESPNSHNEAADDSOLEKANLIELED 539
QY 1565 EGHSGKRGMPHSLSGLQDPIIARMSICSDKSPSECSLIASSPEESWPACOKAYNLNRT 1624
Db 540 DSHSGKRGMPHSLSGLQDPIIARMSICSDKSPSECSLIASSPEESWPACOKAYNLNRT 599
QY 1625 PSTVTLNNTAPTNRANQNPFDEIGRETQSVILRPSPNPTAVQENLKSMAHKRSOR 1684
Db 600 PSTVTLNNTAPTNRANQNPFDEIGRETQSVILRPSPNPTAVQENLKSMAHKRSOR 659
QY 1685 SSVYTRLSKDASELH-AASSESTGFGGERESIL 1715
Db 660 SSVYTRLSKDPPELHAAASSESTGFGGERESIL 691

RESULT 10
Q6MZU2 PRELIMINARY; PRT; 664 AA.
AC Q6MZU2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFp686A19189 (Fragment).
GN Name=DKFp686A19189;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human esophagus tumor;
RA The German Human cDNA Consortium;
RA Bloeker H.; Boecker M.; Mewes H.W.; Weil B.; Amid C.; Osanger A.;
RA Fobo G.; Han M.; Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640878; CAE45935.1; -
DR InterPro; IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON_TER
FT NON_TER
SQ SEQUENCE 664 AA; 72879 MW; 678BA3EF27DA1BDF CRC64;

Query Match 30.9%; Score 2732; DB 2; Length 664;
Best Local Similarity 80.6%; Pred. No. 1.6e-133;
Matches 536; Conservative 31; Mismatches 58; Indels 40; Gaps 3;

QY 1090 PRPPGYSOPASVCSSAFNGFPFGVSPQPHSSYSGLSGPPQHPFYNR----- 1139
Db 1 PRAPGYSOPSPVCSSTSFNGFPAGVSPQPHSSYSGMTGTPQHPFYNRPFPAYLYTP 60
QY 1140 -----AAVPATGSSLLSSMTVDVVCCKLQIQLDON 1171
Db 61 RYYPGSGSHLISRPVSKTSLPRDQNLGSGPAGVPVLNLSNDAVCEKQIQLDON 120
QY 1172 NMMPOYCTTIKKNINGRVLSCQNIDELKKEWAMNFGDMLFRSMVLEKRSVSVQVPDP 1231
Db 121 SMLPOYCTTIKKNINGRVLSCQNIDELKKEWAMNFGDMLFRSTVLEMAESHVVPDP 180
QY 1232 PRFLNENSSAPVPHGSARRSSHTLPLTELSSQTPYTLNFSPEELNTLGLDGAPRHSN 1291
Db 181 PRFLSSSSGPA PHGPARRAASHNELPHTLSSQTPYTLNFSPEELNTLGLDGAPRHSN 240
QY 1292 LSWQSOTRRTPSLSSNQSSSIEISKLTDKVQAEYRDAYREYIAQMSOLEGGTGSSTIS 1351
Db 241 LSWQSOTRRTPSLSSNQSSSIEISKLTDKVQAEYRDAYREYIAQMSOLEGGTGSSTIS 300
```

```
QY 1352 CRSSPHSTYIYICQSSSGGSIHSTLEQERKGEKQEDGRKSFMLKRGDVIDYSSSGVST 1411
RA 1353 CRSSPHSTYIYICQSSSGGSIHSTLEQERKGEKQEDGRKSFMLKRGDVIDYSSSGVST 1411
DB 301 CRSSPHSTYIYICQSSSGGSIHSTLEQERKGEKQEDGRKSFMLKRGDVIDYSSSGVST 360
QY 1412 NEASPLDITTEDEKSDGSKLLPGKSSSRPSLFQTDLKLKGLGRLYQKLPDEDSG 1471
RA 1413 NEASPLDITTEDEKSDGSKLLPGKSSSRPSLFQTDLKLKGLGRLYQKLPDEDSG 1471
DB 361 NDASPLDITTEDEKSDGSKLLPGKSSSRPSLFQTDLKLKGLGRLYQKLPDEDSG 420
QY 1472 TGRVQITPHCSKMTIRKELKAKQREKSCASPFQHSABPIRTFTIKAKYELSDALLDKKSDS 1531
RA 1473 TGRVQITPHCSKMTIRKELKAKQREKSCASPFQHSABPIRTFTIKAKYELSDALLDKKSDS 1531
DB 421 TEESDNTPLKDKDKRKAEGKVERVPKSPHSAEPITFTIKAKYELSDALLDKKSDS 479
QY 1532 GVRNESPNNHLSHNEAADDQLEKANLIELEDEGHSGKRGKMPHSLSGLODPIIARMSIC 1591
RA 1533 GVRNESPNNHLSHNEAADDQLEKANLIELEDEGHSGKRGKMPHSLSGLODPIIARMSIC 1591
DB 480 GVRSESPNNHLSHNEAADDQLEKANLIELEDEGHSGKRGKMPHSLSGLODPIIARMSIC 539
QY 1592 SEDKKSPECSLIASSPESPACOKAYNLARTPTSTVTLNNTAPTNRANQNFDEIEGIR 1651
RA 1593 SEDKKSPECSLIASSPESPACOKAYNLARTPTSTVTLNNTAPTNRANQNFDEIEGIR 1651
DB 540 SEDKKSPECSLIASSPESPACOKAYNLARTPTSTVTLNNTAPTNRANQNFDEIEGIR 599
QY 1652 ETSQVILRPSPNPPTAVONENLKSMAHKRSQSSYTRLSKXDSLH-AASESTGFGEE 1710
RA 1653 ETSQVILRPSPNPPTAVONENLKSMAHKRSQSSYTRLSKXDSLH-AASESTGFGEE 1710
DB 600 ETSQVILRPSPNPPTAVONENLKSMAHKRSQSSYTRLSKXDSLH-AASESTGFGEE 659
QY 1711 RESIL 1715
RA 1712 RESIL 1715
DB 660 RESIL 664
PRT; 1604 AA.
PRELIMINARY;
Q7KVP5 Q7KVP5
AC Q7KVP5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE CG30387-PC.
GN ORFNames=CG30387;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.N., Cawley S., Dailke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,
```

```
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M.P., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Zhong F.N., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195 (2000).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=24242605; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=24242607; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.,
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective."
Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=24242609; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
[5]
RN SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RX FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003453; AAM70867.2; -
DR HSP; Q00420; IAWC.
DR InterPro: IPR002110; ANK.
DR Pfam; PF00023; Ank; 10.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 12.
DR PROSITE; PS00088; ANK_REPEAT; 9.
DR PROSITE; PS02097; ANK_REPEAT_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 1604 AA; 176580 MW; 3259E2F4F183229A CRC64;
Query Match 27.4%; Score 2421.5; DB 2; Length 1604;
Best Local Similarity 35.8%; Pred. No. 8.7e-117;
Matches 584; Conservative 297; Mismatches 585; Indels 165; Gaps 42;
QY 1 MSVLISQSVINYVEENIPALKALLE-KCKDVEDNECGOTPLMAAEQGVNVEIKLLK 59
DB 1 MGSIGHRALLOQYIDNNDISGLRAILDSRHULTIDRDENATTVLMVAVRGTLFAVRBFLA 60
```


QY 1044 DVKTELPCTVNDPKLEIITADYRAAREQINIGLAVPPLPHEGP-----PRPPSGVYQSP 1099
Db 1086 DLRIPLFTINLPYLRKLUKEQOQTIE--DEGSLVIOARPSVNTWRQPAFTTVPSP 1143
QY 1100 ASVCSSAFNGPPGGVVSQPHSSVYSGLSGQHPFYNAAPATCGSSLL---LSSMTV 1156
Db 1144 QAPVPYQMFQNEYPAN---ELSRNLSTSTETVPPLINSFS--DSFGDDILQTKLTDLT 1198
QY 1157 DVVCEKLRQIEGLDQNNMPOYCTTIKANINGRVLSOCNTDELKEMANFMGWHFRSM 1216
Db 1199 EGVISLIDRIEDM-KPALPKLAPVLRENAINGRVLKHCDMPDLKSVLGSFGHWFRL 1257
QY 1217 VLEWRSVE-----SOVVPEDPRELNENSSAPV-----PHGESARS---SHTPEL 1258
Db 1258 ITTLRECELRPRKQOQOQOQGALEAPSNVPMIKQVTDALMQPRESLSKNSVHMEKQ 1317
QY 1259 LT---ELSSQTPYTLNPSFELNTLGLDEGAPRHSNLSWOSQTRTPSLSSL-----N 1308
Db 1318 VTLEEQMICGTLQTLNEEAYE-DVASSERSPTGEMLAQAOLQAPIRESSEFGSPSD 1376
QY 1309 SQSSSIEISKLTQKVQ---AEY-RDAYREYIAQWSQI-----EGGTGSSTI---SGRSS 1355
Db 1377 QKQYGVKISNNNNNNQYLHAEYRNSVSSHLSQSLTLVGAPVGHGGGGSHLHGNGNDL 1436
QY 1356 PHSTY-----YIGQSSGG-----SIHSTLEQER-----GKEGE 1384
Db 1437 SDSTLDMHVDSVFGGGGGGYHRASQISISELHSHESKLDTPSAGAAWTPPLGASGS 1496
QY 1385 LKQEDGRKSFIMKRGDV-----IDYSSSGVSTNEASPLDP-----ITREDEKSPQSGS 1432
Db 1497 VPAPSGRESLTKQGSVKADKRVSIQOMATSNNNNNSTKLTNVEVSRQEVQAGK 1556
QY 1433 KLPLPGKSSRRPS---LFQTDLKLGGGLRYQKLPDSDESGTGRVQIPTHCHSKMIRTKR 1489
Db 1557 RLATKPPGPRPASLIITRDNSSQFOLLRSSVYDDVVAQSHRT-----TIRTTTL 1608
QY 1490 LKAKQRECAPS 1500
Db 1609 LEQOEHEESAP 1619
RESULT 13
Q9W210 PRELIMINARY; PRT; 1678 AA.
AC Q9W210
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE CG30387-PB.
GN ORPNames-CG30387;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Rosler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodsk A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu H.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003453; AAF46711.4; --
DR HSSP; P42773; 1IHB.
DR IntAct; Q9W210; --
DR FlyBase; F5gn0050387; CG30387.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 10.
DR PRINTS; PR01415; ANKYRIN.

Qy 926 LVTFEDWFSDISPOTMRRLNLIVSVTGRLLRANQITFNWDRLASWINLTEQPVRTSTWLIL 985
 :
 Db 1018 VLTDYDFSDVNSRMRLLMNVITVELLKAFQIEFWSYRLSSWINLTEQMLRASMIVL 1077

Qy 986 YLBEE---TEGLPDQWTLKWTYERISKNIPTTKQVEPLLEIDGDIRNFVFSLSTRTPVLVAR 1043
 :
 Db 1078 HHQDQWMSNADESVSLSQSYVEKLRPKLAYLREAAPLLLEDRDRKLDFAFLQLHKHSDLLVA 1137

Qy 1044 DVKTFLPCTVNLPDKRLREIITADVRAREQINIGLAYVPPLPLEGP-----PRPPSGYSQP 1099
 :
 Db 1138 DLRIFLFTTINLDPYLRKVLEKBOQTIE--DEGLVIQARPSVSNWRQFPAPTYVYPSP 1195

Qy 1100 ASVCSSASFNGPFPGGVSPQPHSSYSYGISGCPHFYNRAAYPATGSSL----LSSMTV 1156
 :
 Db 1196 QAVPPYQMONEYPAN----ELSRNLSLTSTEPVTPLINSPS-DSFGDDILQTKLDTLTV 1250

Qy 1157 DVVCEKLRQIEGLDONMPQCYTIKANINGRVLSQCNIDELKKEMANNFGDWHUFRSM 1216
 :
 Db 1251 EGVISLDRDIEDM-KPALPKLAPVLRENAINGRVLKHCDMPDLKSLVGLSFGHWELFRLL 1309

Qy 1217 VLEMRSVE-----SOVVPEDPRLFNNSSAPV-----PHGESARES---SHTELP 1258
 :
 Db 1310 ITTLRECELRPRKQQQQQPGALEAPSNVPMIKVD TALMQPPRESLSRKNSVSHMEKQ 1369

Qy 1259 LT---ELSSQPTYTLNPFSEELNTGLDGCAHRHSNLSWOSQTRRTPTSLSSL-----N 1308
 :
 Db 1370 VTLEEOMICGTQLTNEEAVE-DVASERPSPTEGM LAVAQLOLAPIRESSEFGSPSDD 1428

Qy 1309 SQDSSTEISKLTKVKQ---AEY-RDAYREVIAQMSQI-----EGTGGSTI---SGRSS 1355
 :
 Db 1429 KQGYGVKISNNNNNNQYLHAEYNRSVSSHLSQSLTLVGAPVGHGSGGSHLHGNGNDL 1488

Qy 1356 PHSTY-----YIGQSSSG-----SIHSTLQEOR-----GKEGE 1384
 :
 Db 1489 SDSTLDLMHVDSVFGGGGGGYHRASQGISSELLHESKLDTPSAGAAWTTPLLGA SGS 1548

Qy 1385 LKQEDGRKGFPMKRGDV-----IDYSSSGYSTNEASPLDP-----ITEEDEKSQDSGS 1432
 :
 Db 1549 VPAPSGRESILTKOOGSVKADKRVSIQMATSNNNNNNKLTNPVENVESREQVOGAGK 1608

Qy 1433 KLPLGKKSSERP S-----LFQTDCLKLGGLRYKQLPSDEDESGTGRVQIPTHCSKMIRTXR 1489
 :
 Db 1609 RLTTKPPGPRPASLIITRNDSNSQPQLLRSSSDYDDVEAQEHRT-----TIRRTL 1660

Qy 1490 LKAQRKECASP 1500
 | : : | : : |
 Db 1661 LEQQEBSAP 1671

RESULT 14
 Q9H9E4 PRELIMINARY; PRT; 543 AA.

ID Q9H9E4 AC Q9H9E4 PRELIMINARY; PRT; 543 AA.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ12811.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14702039; DOI=10.1038/ngl1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi K., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayangi T., Wagatsuma M., Shiraori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamiyama K., Katsuma N., Sato K., Tanikawa M., Yamazaki M.,

Db 542 YFTDILLKFLDR--YDLEWYSLNYGLSRKLRLLQLVAFCHPPG--POSDDQPMFVR 598
Qy 598 FLFTDYNRSLSGVGETSLAEMIATLSDACEREFGLATRLFRVERTE--ESQKKKKKKT 655
Db 599 FHFAEAG--AAPGDAVALMLASLFDIAEHAHYSGLATGLYRAFPLKATGGKWKRM 657
Qy 656 CCLPSFVIFLPIVGCIIAGITLLAIFRVPDKHLTVNA---ILISIASVVGAP---VLNC 709
Db 658 CCMF--VVLMEFELGML--GILATASLSIVSYSEYGLGREGIEAIAIYVLLGILLAGAIANL 713
Qy 710 RTWQVLDLSLNSQKRLHSAASKLHLKLSGPFMKVLKCEVELMARMAKTIDSFTQNTQR 769
Db 714 HAWSKLIGALFMSOGKHLKRAFNS-----NEAAPTALGAEVSMLTDMVRCILDAFTGQOSR 769
Qy 770 LWIIDGLDACQDKVQLMDTVRVLPFS--KGPFAIFASDPHIIKAINONLNSVLKDS 827
Db 770 LVGVVDALDSCDTERLTILNAVQTLTLLSGPQPPFVLLAVDPHVIKAAEANSRRLFTTEG 829
Qy 828 NINGHDYMRNIVLHPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTD----- 879
Db 830 GIGGHDFLURNLVHLPVYLQNSGLRKVQRAQNTAMSAVRRAMPDLS---RDDQPHLGHSA 886
Qy 880 --RRVSQNS--LGEMTKLGSKTALNRRDITYRRQWQ-----RTITRQMSFDLYKL 925
Db 887 SVRLSNASEIMSQEKLAMPSSSRGSKKLRVSDSIASSIGNLHKLQGNPPVDLSKI 946
Qy 926 LVTEDFWSDISPOTRRLNLIVSVTGRLLRANOITFNWDRLASWINLTQOWPYRTSMIL 985
Db 947 ILTDYFSDVNPMSRRRLNVIYITVRLKAFQIDFSWYRLSSWINLTQOWPLRASMIIV- 1005
Qy 986 YLEETEG---LPDQMTLKTMYERISKNIPTTKDVEPLIEIDGDIRNFEVFLSSRTPVLVA 1042
Db 1006 -LEHDQGDSPDSDSVLSQAVDYKVRPKLTCLEAAALDLDLDRDKLDAFLQLHKSLLV 1064
Qy 1043 RDVTFPLCTVNLDPKUREIIVADVRAAREQINIGGLAYPPLPLHEGPPRPPSGYSQPASV 1102
Db 1065 SDRIFLFPFTNLDPYLRKVLKEDQALE--DEGVYGTNALMGEPLPKP----- 1112
Qy 1103 CSSASFNGPPPGGVVSPQPHSSYVYSLGSPQHPFVNRAAVPATG----- 1146
Db 1113 ---GSVGG---GPIITEHPHEQHSITLGSNS-----GRSKTTTALOKSSNGLATSPPI 1162
Qy 1147 ----SLLLSMTVDVCEKLRQIEGLDQNNMPQYCTTIKKANINGRVLSQCNIDELKKE 1202
Db 1163 DIDLNSQLSSLTVEQLIELEKVNDL-KPAMERTAPILRENAISGRVLMYCNLEELKSV 1221
Qy 1203 MAMNFGDWHLPRSMVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTPLPTE- 1261
Db 1222 LRLSFGHWMFKLLVLTALR--EASVTPASRKLSTKIT-----SPAKGTNDSVENQEP 1272
Qy 1262 LSSQTPYTLNFSFBEELNTLGLDEGAPRHSNLSWQSOTRTPSLSLNSQDSSIEISKLTD 1321
Db 1273 IAQSTP-----PFGSSTS-SFQPIRQKSNLLEKQ-VTLEEQMICG 1311
Qy 1322 KVQAEYRDVAYREYIAQMSQLEGGTSGSTISGRSSPHSTY-----IGQSSSGS 1370
Db 1312 ALQTLNEDAFEDVYSS-----SERPFTGEMFQYLAPIRESSEIGS 1353

Search completed: February 5, 2005, 19:33:57
Job time : 198.5 secs

This Page Blank (uspto)